View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Dundee Online Publications

University of Dundee

Mining Ethos in Political Debate

Duthie, Rory; Budzynska, Katarzyna; Reed, Chris

Published in:
Computational Models of Argument

DOI:
10.3233/978-1-61499-686-6-299

Publication date:
2016

Document Version
Final published version

Link to publication in Discovery Research Portal

Citation for published version (APA):
Duthie, R., Budzynska, K., & Reed, C. (2016). Mining Ethos in Political Debate. In P. Baroni, T. F. Gordon, T. Scheffler, & M. Stede (Eds.), Computational Models of Argument: Proceedings from the Sixth International Conference on Computational Models of Argument (COMMA). (Vol. 287, pp. 299-310). (Frontiers in Artificial Intelligence and Applications; Vol. 287). Netherlands: IOS Press. DOI: 10.3233/978-1-61499-686-6-299

General rights
Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain.
• You may freely distribute the URL identifying the publication in the public portal.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 28. Apr. 2017

Computational Models of Argument
P. Baroni et al. (Eds.)

© 2016 The authors and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/978-1-61499-686-6-299

Mining Ethos in Political Debate

Rory DUTHIE [a], Katarzyna BUDZYNSKA [a,b], and Chris REED [a]
[a] *Centre for Argument Technology, University of Dundee, UK*
[b] *Institute of Philosophy and Sociology, Polish Academy of Sciences, Poland*

Abstract. Despite the fact it has been recognised since Aristotle that ethos and credibility play a critical role in many types of communication, these facts are rarely studied in linguistically oriented AI which has enjoyed such success in processing complex features as sentiment, opinion, and most recently arguments. This paper shows how a text analysis pipeline of structural and statistical approaches to natural language processing (NLP) can be deployed to tackle ethos by mining linguistic resources from the political domain. We summarise a coding scheme for annotating ethotic expressions; present the first openly available corpus to support further, comparative research in the area; and report results from a system for automatically recognising the presence and polarity of ethotic expressions. Finally, we hypothesise that in the political sphere, ethos analytics – including recognising who trusts whom and who is attacking whose reputation – might act as a powerful toolset for understanding and even anticipating the dynamics of governments. By exploring several examples of correspondence between ethos analytics in political discourse and major events and dynamics in the political landscape, we uncover tantalising evidence in support of this hypothesis.

Keywords. Character of speakers; Ethos attack; Ethos support; Natural Language Processing; Parliamentary debates; Sentiment analysis

1. Introduction

Ethos is defined as the character of the speaker [1], i.e. the character of the person who is the participant of communication. It has been extensively studied in disciplines such as rhetoric, epistemology and social psychology for the major role it plays in communication and society.[1] Ethos forms a crucial part of a debate along with two other means of persuasion: pathos which is the audience's emotions; and logos which is the use of reasoning. In [29], arguments containing ethos (argument from expert opinion) are studied from a logos perspective, however pure ethos has been studied less so.

This paper aims to demonstrate that linguistically oriented AI, by making use of a large amount of data, can offer insights and improve our understanding of how ethos influences the interaction between communicating agents and the formation of social structures. In the political sphere, knowing who supports whose ethos (see **Ex. 1**)[2]; who attacks whom (**Ex. 2**); whether the sentiment is mutual (e.g. Radice used to attack Pawsey and vice versa in the example); which political party the person represents (e.g. Patten supported Ewing twice even though they are from opposite parties); is a powerful tool

[1] Although trust attracted quite a lot of attention in AI (cf. [3,20,22]), this notion is used differently than the notion of ethos presented in this paper.
[2] Examples are taken from UK parliament of 1979-1990.

for understanding the dynamics of governments such as the creation of cliques and coalitions, and the rise and fall of rebellious behaviour. However, manual analysis of such large data-sets in broadcast and social media, or parliamentary records is a very labour intensive task. In this paper, we propose a method of automating such analysis.

Example 1 Mr. Chris Patten said, *The hon. Member for Falkirk, East (Mr. Ewing) in his admirable speech, put the position much more clearly than I could.*

Example 2 Mr. Giles Radice said, *In doing so he (Mr. Pawsey) failed to face up to his responsibility both to the House and to the schools of England, Scotland and Wales.*

We use a pipeline of natural language processing techniques to extract the information from the linguistic surface of diplomatic language in expressing opinions during UK parliamentary debates. For example, the phrase, "admirable speech" in **Ex. 1** can suggest support for Mr. Ewing's ethos,[3] while "failed (...) to his responsibility" can be used as a cue that Mr. Pawsey was attacked. This task requires several challenges to be addressed. For example, the dialogical context encourages the use of pronouns (see "he" in **Ex. 2**); reported speech (see **Ex. 3**) includes references to other people which are ethotically neutral; or some phrases, which seem positive such as "honorable", are in fact a part of political etiquette. Thus, a system we propose is a pipeline of components that deal with these challenges step by step.

Example 3 Mr. Giles Radice said, *The hon. Member for Rugby and Kenilworth (Mr. Pawsey) said that in the United States and Australia this was a local decision.*

Specifically, the contribution of the paper includes: (*a*) the first freely accessible corpus specifically annotated with tags allowing for the representation of ethotic linguistic structures; (*b*) a system for ethos mining consisting of existing methods such as Part-of-Speech tagging and SVM-based sentiment classifier as well as new techniques such as anaphora resolution, rule-based expression recognition and a reported speech filter; (*c*) software for visualisation of the relationships between politicians allowing the analysis to produce insights into data not normally seen in the political science literature; (*d*) exploratory applications of these visualisation and ethos analytics tools to periods in the historical parliamentary record associated with major political upheaval, demonstrating how the changing political landscape is reflected in, signals in the ethotic interactions in the text.

2. Corpus

Our data is taken from the UK parliamentary record, Hansard, which is an online archive of transcripts of all House of Commons and House of Lords debates dating back to the 1800s (freely available at `http://hansard.millbanksystems.com/`). The archive is organised by the day divided into a number of sessions on different topics. Each turn in the debate consists of the identification of Members of Parliament, MP, followed by their constituency (if this is the first time they have spoken) and their speech.

Corpus	Sessions	Words	Segments	Speakers	Location
Train	30	40,939	387	127	`http://arg.tech/Ethan3Train`
Test	30	29,178	352	126	`http://arg.tech/Ethan3Test`
TOTAL	60	70,117	739	253	

Table 1. Summary of the language resources in the EtHan_Thatcher_3 corpus for mining ethos in Hansard.

2.1. Data

The corpus EtHan_Thatcher_3[4] (see **Table. 1**) was constructed by taking a random subsample of Hansard according to the following rubric: select the first two House of Commons debates over 700 words in length from the day closest to the date(s) at the midpoint(s) of the largest uninterrupted date range(s) (initially the midpoint in the range 4th May 1979 and 22nd November 1990 - viz., 11th February 1985; then at the midpoints between 4th May 1979 and 11th February 1985, and between 11th February 1985 and 22nd November 1990, etc.). This avoids bias for annotators and yielded 60 transcripts, the data in each of which was then cleaned such that any titles and section markers were removed to leave only the speakers, organisations or other entities and the statements they made. The transcripts were then split evenly to give a training set and a testing set. The training set formed the training data for the sentiment polarity classifier and was used as the basis for developing domain specific rules for recognising ethotic sentiment expressions.

2.2. Annotation

The annotation was performed by applying four tags (see **Table. 2** for their frequency) according to the following guidelines:

Source-person. Source tag is used to mark a person who utters the statement.

Target-person. Target is a person who is described by the statement.

Ethos support. Ethos support should be identified when: (*a*) the statement makes explicit mentions of a person, organisation or other entity (excluding groups and assemblages) except when this is reported speech; and (*b*) it takes the form of supporting a person's credibility or looking to put them in a positive frame through character supports or supports of work; and (*c*) a support to a person's own ethos should not be analysed as this is deemed to be a fallacy [2]. Compare **Ex. 1**.

Ethos attack. Ethos attack should be identified when: (*a*) the statement makes explicit mentions of a person, organisation or other entity (excluding groups and assemblages) except when this is reported speech; and (*b*) it takes the form of attacking a person's credibility or looking to put them into a negative frame; or (*c*) it may take the form of trying to unbalance authority on a subject giving the attacker more of a right to talk about the subject. Compare **Ex. 2**.

The statements in which speakers refer to other persons are called **Ethotic Sentiment Expressions, ESEs** and the statements which do not contain reference to oth-

[3]Though such a sentiment can in principle be cancelled or reversed by subsequent linguistic material, in practice in our corpus such situations almost never occur.

[4]The corpus is named as so due to the annotation of session transcripts at different time periods. EtHan_Thatcher_1 containing an original 30 sessions which was extended to EtHan_Thatcher_3 and EtHan_Thatcher_2 containing a subset of EtHan_Thatcher_3 for agreement calculations.

Source-person	243
Target-person	212
Ethos support	179
Ethos attack	560
TOTAL	1,194

Table 2. Occurrences of tags in EtHan_Thatcher_3

ers are denoted as **non-ESEs**. The polarity of these statements is then expressed by the use of abbreviation **+ESE** for positive sentiment (ethos support) and **-ESE** for negative sentiment (ethos attack). The data was analysed according to the standard of argument representation, i.e. Argument Interchange Format (AIF) [24], using the OVA+ annotation tool [9] (freely available at `http://ova.arg-tech.org`) and stored in the AIFdb database [10] (`http://aifdb.org`). Annotation is below sentence level but above word level.[5]

2.3. Evaluation

In order to evaluate annotation, we selected a subset of data used in the EtHan_Thatcher_3 corpus. The selection followed the same method as applied to the whole dataset. The total size of this subset comprises 10% of the EtHan_Thatcher_3 corpus with 6 sessions containing 7,267 words, 91 segments and 30 speakers. Cohen's kappa for recognising whether the statement is ESE or not gave the value of $\kappa = 0.67$. For ethotic statements, $\kappa = 0.95$, when it is a support or an attack. For source-person of an ethotic statement, $\kappa = 1$ and for target-person it was $\kappa = 0.84$, all for two coders.

3. Automation

3.1. System Architecture

The architecture of the software system for mining ethos consists of three stages, five layers and eight components (see **Fig. 1**). The three stages consist of the ESE / Non-ESE stage, the +/- ESE stage and the network stage. The ESE / Non-ESE stage takes an input of cleaned text transcripts from the EtHan_Thatcher_3 test sub-corpus and classifies each segment as either an ESE or non-ESE. The +/- ESE stage then gives the polarity of ESEs, ESEs with positive sentiment (corresponding to ethos support, as in **Ex. 1**), and ESEs with negative sentiment (corresponding to ethos attack, as in **Ex. 2**). Finally, the network stage provides a visualisation of all ESEs as edges between each participant in the debate.

In the ESE / Non-ESE stage, there are three layers consisting of five components. The parsing layer uses plain text from the EtHan_Thatcher_3 test sub-corpus and applies three different methods to it: Named Entity Recognition (NER), Part-Of-Speech (POS) tagging and a set of domain specific rules. The output is Agent Reference Expressions (AREs) which are any statements referring to another person, organisation or agentive entity. Given the dialogical nature of the material, many statements do not refer to the target-person by their name explicitly, but e.g. by a pronoun (see "he" in **Ex. 1**), by

[5] The annotation is visualised as directed graphs where support is marked as Default Inference, attack as Default Conflict, source-person is in the node with the statement, and target-person – in the node which refers to ethos.

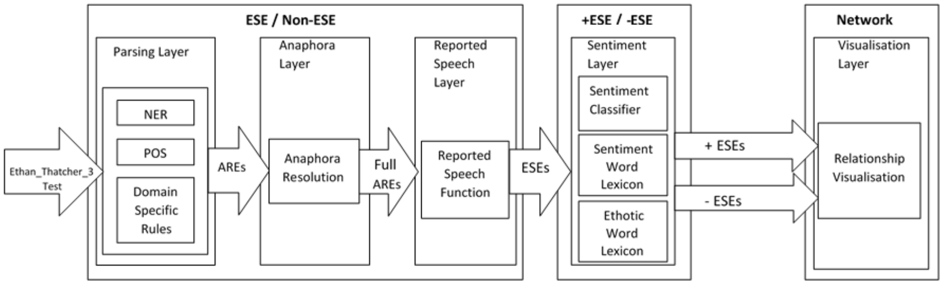

Figure 1. A text analysis pipeline for ethos mining: the extraction, polarisation and networking of ESEs from Hansard sessions in plain text transcripts.

a region MP represents (see "The hon. Member for Falkirk, East" in **Ex. 2**) or by a functional role e.g. "the Prime Minister". Thus, AREs are then passed to the anaphora layer where both source-person and target-person of the statement are retrieved from the original text. The next challenge is that the repetitions of what has been previously said can be ethotically neutral, especially when an MP wants to remind some thread of the debate which happened many turns earlier (see **Ex. 3**). Therefore, full AREs are passed to the reported speech layer where an ARE is removed if it is not an ethotic expression but a reported speech.

In the +/- ESE stage there is one layer, the sentiment layer, containing two components, the sentiment classifier and the word lexicons. The sentiment classifier and word lexicon components combine to classify ESEs as positive and negative. These two sets are then passed to the Network stage where the visualisation layer displays relationships between people, organisations and other entities. Not attached to **Fig. 1** is the training sub-corpus which is used just for defining domain specific rules and the lexicon for the +/- ESE stage for the sentiment classifier. The techniques of domain specific rules, anaphora resolution, reported speech function and relationship visualisation were developed specifically for the tasks of ethos mining in political debate, and the method of sentiment classification was extended with the development of a lexicon to account for the characteristics of the domain.

3.2. Methods

The ethos mining tool applies existing NLP methods such as Part-of-Speech tagging and an SVM-based sentiment classifier with an existing sentiment word lexicon, and new techniques such as anaphora resolution, rule-based expression recognition, a reported speech filter and an ethotic word lexicon.

Named Entity Recognition (NER). NER, using the Stanford Named Entity Recognizer (Stanford NER) [6] with 92.28% accuracy on the CMU Seminar Announcements information extraction dataset, is performed to extract statements which contain names, organisations and locations from the plain text on the assumption that any specific statement referring to a named entity can in fact be a form of ethotic statement. This is applied to the original text from EtHan_Thatcher_3 test sub-corpus and produces a set of AREs on the assumption that any specific statement made to a named entity can in fact be a form of ethotic statement.

Part-of-Speech (POS) Tagging. POS tagging, using the Stanford POS Tagger [28] with an accuracy of 97.24%, is applied to extract statements which contain pronouns to account for situations such as in **Ex. 2**. This was applied to the EtHan_Thatcher_3 test sub-corpus and then run against the list of already extracted AREs from the NER to account for any duplicate segments extending the list of AREs.

Domain Specific Rules (DSR). We developed rule-based expression recognition to account for the specific language of the political domain. In the House of Commons, the speaker is not allowed to refer to any other MP by name, but by phrases such as "Honourable Gentleman" or "Honourable Lady". The constituency name of an MP can be used in the same respect to address an MP such as in **Ex. 1** "The hon. Member for Falkirk, East" Organisations can also be mentioned under a different name, e.g. "the Government" will refer to the party in charge of the government at that time, and "the Opposition" – to the current official opposition. These rules are then extended with the creation of a list of ethotic words to determine if ethos is held in a particular ARE. A list of 326 ethotic statements were compiled from the EtHan_Thatcher_3 training sub-corpus, containing some words not normally used in day-to-day conversation such as "penny-pinching" and "gerrymandering". These words are common with ethotic attacks. Again the new AREs produced from this component are checked against the list of already extracted AREs to remove duplicates.

Anaphora Resolution (AnaR). We developed this rule-based module with manually defined rules to reconstruct all sources and targets in each AREs. For the source-person, the reconstruction is needed, when a sentence is not the first one in a turn in the dialogue (a turn corresponds to a paragraph in the transcript). In such cases, first the system associates a sentence with a paragraph. Since paragraphs are assigned a source-person, thus this person becomes a source for the sentence. For a target-person, there are two possible cases. First, when the anaphora occurs in situations such as "MP_1 said *MP_2 did this and he did that*", NER technique is used. In the case of sentences such as "MP_1 said *he did this*", the system tracks back to the beginning of the paragraph. If nothing is found, then it looks for the speaker of the previous turn.

Reported Speech Function (RSF). We developed a reported speech filter which aims to remove segments containing neutral reports of what previously has been said by other speakers (thus no ethotic sentiment). The technique uses lexical cues such as "says", "you say" and "told me", and any segment containing these words is removed from the list of AREs. RSF produces a list of ESEs which are then passed to the sentiment classifier.

Sentiment Classifier (SVM, NB, ME). To perform sentiment analysis three machine learning algorithms were considered: Support Vector Machines (SVM), Naïve Bayes (NB) and Maximum Entropy (ME). A C-Support Vector Classification (C-SVC) algorithm [4] from the LIBSVM library was used to classify ESEs into two sets: positive and negative. To perform NB and ME the Stanford classifier library [13] was used. In selecting these methods, we followed the conclusion formulated in [19] that the discourse approach in sentiment analysis is not satisfactory and that supervised learning techniques are needed (which is demonstrated in [21] with the good performance of SVC of 83%). The lexicon (defined in **Section. 3.2**) was passed to the Stanford CoreNLP [14] library in order to perform lemmatization, allowing the frequency of words in the lexicon to be more accurately calculated.

Lexicon (SWL, EWL). To perform sentiment analysis one existing lexicon was used, the sentiment word lexicon (SWL) [8], and one lexicon created, an ethotic word lexicon (EWL). The SWL, contains 2,006 words tagged as positive and 4,738 words tagged as

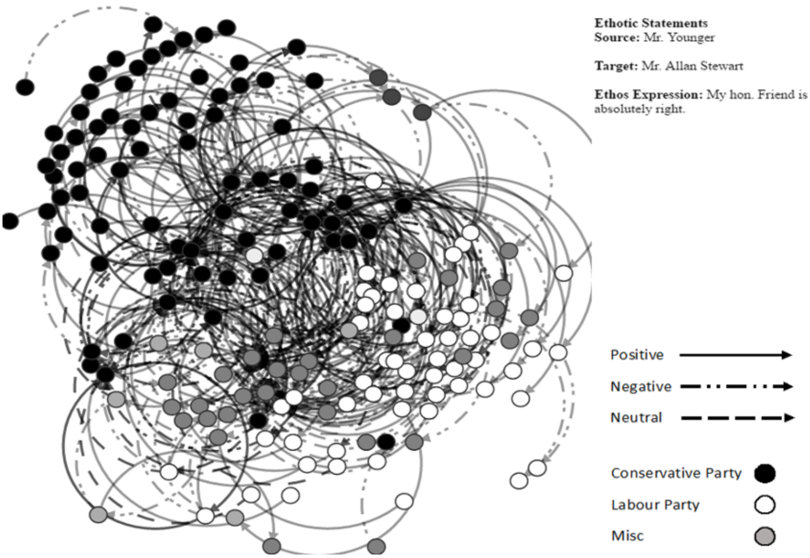

Figure 2. The component of Relationship Visualisation for EtHan_Thatcher_3 test corpus, showing the network of positive and negative relationships in parliament (Available at: `http://arg.tech/EthanVis`).

negative. The EWL is a set of keywords developed using the EtHan_Thatcher_3 training sub-corpus containing 381 tagged sentences with 96 positive and 285 negative from which unigrams, bigrams and trigrams were extracted. Despite the relatively small volume of this set, its advantage lies in its adaptation to sentiment related specifically to ethos in political debate. The removal of non sentiment bearing words and named entities, and the use of n-grams gave 32,858 features overall to be used as training data for machine learning.

Relationship Visualisation. Extracted ESEs with polarity and source and target person were used for visualisation purposes. D3.js a javascript graph visualisation library (available at: `http://d3js.org/`) was used to create force-directed graphs representing positive (coloured as green) and negative (coloured as red) relationships amongst the politicians (see **Fig. 2**). Each edge representing a relationship is associated with a set of ESEs depending on the polarity of the ESE. People are visualised as nodes coloured according to their political party. Nodes are then clustered by political party using a multi-foci technique: nodes which are pulled closer together show that there were either many attacks or supports between them.

3.3. System Results

Results are given for the two stages of ethos mining shown in **Fig. 1**, the ESE / Non-ESE stage and the +/- ESE stage. A result is also given for the combination of these stages. **Table. 3** gives the results of precision, recall and F-score for the classification of ESEs as an ESE or Non-ESE. ESEs are defined as correct if the text they hold contains the corresponding segment in the manual analysis, which is true for exact matches and ESEs holding more text than has been manually annotated.

In **Table. 3** we consider a baseline classifier which predicts only the target class (ESE), common machine learning algorithms (ME, NB and SVM) and the ESE / Non-

ESE/Non-ESE	Precision	Recall	F-score
Baseline	0.29	1	0.45
SVM	0.30	0.30	0.30
NB	0.20	0.94	0.32
ME	0.46	0.27	0.34
NER, POS, DSR, AnaR, RSF	0.62	0.77	0.69*
POS, DSR, AnaR, RSF	0.64	0.76	0.70*

Table 3. Results of automatic extraction of ESEs from EtHan_Thatcher_3 Test corpus. We report precision, recall and F-score for classifying ESEs as ESE and Non-ESE. The star symbol (*) denotes the classifier above the baseline F-score.

+/- ESE	Precision	Recall	F-score
Baseline	0.50	1	0.67
NB, SWL	0.58	0.57	0.57
ME, SWL	0.6	0.65	0.62
SVM, SWL	0.64	0.59	0.62
NB, SWL, EWL	0.74	0.67	0.71*
ME, SWL, EWL	0.71	0.73	0.72*
SVM, SWL, EWL	0.78	0.78	0.78*

Table 4. Results for the sentiment classifier based on a Macro-average of results of both positive and negative classifications. We report precision, recall and F-score for a baseline classifier and machine learning classifiers. The star symbol (*) denotes the classifier above the baseline F-score.

ESE stage of our system, containing NER and with NER removed. Of these algorithms both of our systems perform above the baseline F-score by 53%. To identify people within Hansard it would be logical to perform NER to extract names from text. Although this would be true for most cases of dialogue, due to UK parliamentary rules, the number of instances where names are used explicitly are few. This can cause the problem of many false positives being extracted by the ESE / Non-ESE stage. In removing NER we observe an increase in precision on the ethos mining system with only a slight drop in recall.

In **Table. 4** the results of +/- ESE classification are reported with comparison of common machine learning techniques to a baseline classifier with a macro-averaged precision, recall and F-score of the majority class (negative) and the minority class (positive). Comparison is made two different training lexicon, SWL and EWL in **Section. 3.2**. The results indicate that known ethotic words which we developed for the EWL are crucial in obtaining high F-score on sentiment classification of ESEs. Using the same set of features an SVM classifier outperforms both a Naïve Bayes Classifier and a Maximum Entropy Classifier with an overall $F1$-score 16% above the baseline.

ESE/Non-ESE & +/- ESE	Precision	Recall	F-score
Baseline	0.14	1	0.25
Full System	0.55	0.65	0.60*

Table 5. Results are provided for the combination of the ESE / Non-ESE stage and the +/- ESE stage.

In **Table. 5** the results of the combination of the ESE/Non-ESE stage and the +/- ESE stage are given. A true value is only given when the system correctly identifies an ESE and gives the correct sentiment polarity, when compared to manual analysis. A drop in overall F-score from **Table. 3** is observed due to the error margin, reported in **Table. 4**. However, when calculating the baseline for the full system this gives F-score 0.25, putting the full system, containing the ESE / Non-ESE stage and SVM +/- ESE stage, 40% ahead of the baseline.

4. Scaling up

In this section, we explore two examples of correspondence between ethos analytics in parliamentary discourse and major political events In other words, we do not aim here to evaluate the ethos mining tool, but to illustrate its analytical potential by comparing the output of the automatic system not to manual annotation, but to political science publications and news articles from the considered time periods.

February 1st 1997 to April 30th 1997, 53 text transcripts focusing on the final stages of the Conservative government before Labour leader Tony Blair became Prime Minister.[6] In this time, it was documented that John Major, the then Prime Minister, was struggling to keep his own party on side [7]. This is evident in the analysis with eight ethotic attacks coming from his own party and two attacks coming from Tony Blair, the leader of the opposition at the time, where the average number of attacks is two. Following the loss of the general election to the Labour party a new leader of the Conservatives was elected. Interestingly, in the lead up to the general election, the proposed candidates for the Conservative Leadership election are more prominent in the visualisation as seen in **Table. 6** where the mean for number of supports and attacks for a politician is two. Many supports and attacks of the potential leaders hint at their impending desire to run for party leadership as a high number of either show that the potential leaders are more prominent in debate.

Potential Conservative Leaders	Supports	Attacks
William Hague	33	30
Ian Lang	17	20
Stephen Dorrell	22	10
Michael Howard	4	4
Peter Lilley	3	0
John Redwood	2	0
Kenneth Clarke	0	0
MEAN AVERAGE	2	2

Table 6. Supports and Attacks on ethos of Conservative Leader proposed candidates.

November 30th 1978 to January 20th 1979, 32 text transcripts focusing on a period of time in the UK known as the Winter of Discontent. In this period there were multiple strikes by workers in the UK, putting pressure on the then Labour Prime Minister James Callaghan [27]. This period was characterised by two significant changes in the

[6]Note that ethos analytics was run on a larger set than the EtHan_Thatcher corpus, because we used all transcripts from a given analysed period.

political landscape: first, the growth of mass infighting in the Labour party; and second, Margaret Thatcher becoming Prime Minister on May 4th 1979. These political dynamics are reflected in two ways in the ethos analytics. The Prime Minister James Callaghan had a total of eleven attacks on his ethos, where the mean across all MPs is one. Half of these were from Labour party members, reflecting the deep discontent at his leadership. The infighting which followed is also reflected by the ethos analytics. Shirley Williams, a Labour member at the time, has for example a total of eight attacks on her ethos and seventeen supports (only two come from other Labour members). In the years following the general election, and after the loss of Williams' seat, she became a founding member of the Social Democrat Party (SDP) [5].

5. Related work

Although, as far as we are aware, automatically extracting linguistic expressions of ethos has not previously been explored, ethos mining builds on methods and techniques developed for sentiment analysis and argument mining. The closest approaches to ours include the application of NLP techniques to the UK Hansard to build a database of claims associated with their parliamentary authors [18, 19]; the use of a lexicon based and classification approach in analysing sentiment of UK parliamentary debates [25]; and mining of arguments from the Canadian Hansard parliamentary record [17]. It is however, important to note that these works perform different tasks to ethos mining so the results are not directly comparable.

Sentiment analysis is the classification of documents, sentences or individual words as either positive or negative. Sentiment classification using machine learning can achieve over 80% accuracy [21] when performed on large feature vector sets using only unigrams as features. [12] describes feature-based sentiment analysis using a lexicon of sentiment bearing words to classify text. In [18], a system was developed to extract politicians' statements on specific topics in order to increase the accuracy of queries in UK Hansard. To do this, NLP techniques such as NER and POS tagging were used giving a satisfaction rating of 32% on an ordinal scale. In [19] the approach was extended by applying discourse sentiment analysis, with an accuracy of 44%. In [25], NLP techniques such as POS tagging were applied to parliamentary debates to obtain features for machine learning classifiers. These were then compared to two lexicon based sentiment approaches, an off-the-shelf lexicon approach, SentiWordNet 3.0 and a domain specific lexicon approach. When compared, the machine learning classifiers out performed the lexicon based sentiment approaches, with an accuracy of 61.75%.

Argument mining (also called argumentation mining, see e.g. [16, 23, 26, 30] for an overview) is the automatic extraction of argument from text over many different domains. In [15], legal text is broken down into sentences which then have features extracted. Sentences are then classified as either argumentative or non-argumentative with an accuracy of 68% for legal texts. In [11], claim detection is explored using NLP techniques to extract features of argument. Using an SVM with data extracted using parsing and POS tagging, a precision of 9.8 and a recall of 58.7 were achieved. In [17], a corpus of 138 sentences from Gay Marriage political debates was annotated by three coders with an inter-annotator agreement using weighted kappa of 0.54 for stance (a users stance on a sentence) and 0.46 for frames (pre-existing arguments which highlight an aspect of an argument), with 90% agreement on statements between at least two of three annotators.

An SVM classifier was trained using a bag of words approach, distributed word representations of stance and frames with similarity calculations and the stance of each statement, either pro- or con-, as a feature. Frames were then identified in political speeches with an overall accuracy of 68.9%.

6. Conclusion

Whilst ethos is well-recognised as a critical, load-bearing component in successful communication, it has attracted relatively little attention in AI, particularly with respect to the way in which it is made manifest in language. We have presented the first systematic treatment of ethos from a linguistically-oriented AI perspective, including a simple coding scheme applied to the UK parliamentary record, Hansard, resulting in an annotated corpus which is openly available. We have shown that a text analysis pipeline of hand-crafted domain specific rules, structural linguistic methods and supervised learning techniques, improved by our lexicon of ethotic words, can deliver strong performance on identifying expressions of ethos, when compared to related work in argument mining of political debates. It is important to note that results achieved in argument mining cannot give a definitive comparison due to the difference in logos and ethos. By aggregating the results into visualisation and analytics, it becomes possible to identify patterns in new, unannotated datasets. Indicative and exploratory analysis of historical records suggests that major events and trends in the political landscape are reflected in, or anticipated by, the ethos analytics of the parliamentary record. On the one hand, this opens up an exciting new research programme to understand the relationships between ethos-oriented linguistic interactions amongst politicians and the historical events with which they are associated; but on the other it also raises the intriguing possibility that such techniques may be able to link contemporary ethos dynamics with the political events they presage.

Acknowledgements

This research was supported in part by EPSRC in the UK under grant EP/M506497/1 and in part by the Polish National Science Centre under grant 2015/18/M/HS1/00620. We also thank Marcin Koszowy for providing second annotator analysis.

References

[1] Aristotle. *On Rhetoric (G. A. Kennedy, Trans.).* New York: Oxford University Press., 1991.

[2] K. Budzynska. Circularity in ethotic structures. *Synthese*, 190:3185–3207, 2012.

[3] Chris Burnett, Timothy J. Norman, and Katia Sycara. Trust Decision-Making in Multi-Agent Systems. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, IJCAI*, 2011.

[4] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[5] Liberal Democrats. Shirley Williams, 2016. `http://www.libdems.org.uk/shirley_williams` [Last Accessed: 02/02/16].

[6] Jenny Rose Finkel, Trond Grenager, and Christopher Manning. Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. *Proceedings of the 43nd Annual Meeting of the Association for Computational Linguistics (ACL, 2005)*, pages 363–370, 2005.

[7] Gov.uk. History of Sir John Major - GOV.UK, 2016. `https://www.gov.uk/government/history/past-prime-ministers/john-major`[Last Accessed: 02/02/16].

[8] M. Hu and B. Liu. Mining and Summarising Customer Reviews. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, Washington, USA*, 2004.

[9] M. Janier, J. Lawrence, and C. Reed. OVA+: An argument analysis interface. In *Computational Models of Argument (COMMA)*, 2014.

[10] J. Lawrence, M. Janier, and C. Reed. Working with Open Argument Corpora. In *European Conference on Argumentation (ECA)*, 2015.

[11] Marco. Lippi and Paolo. Torroni. Context-Independent Claim Detection for Argument Mining. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI*, 2015.

[12] Bing Liu. Sentiment Analysis and Subjectivity. *Handbook of natural language processing*, 2:627–666, 2010.

[13] Christopher Manning and Dan Klein. Optimization, maxent models, and conditional estimation without magic. In *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology: Tutorials-Volume 5*, pages 8–8. Association for Computational Linguistics, 2003.

[14] Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 55–60, 2014.

[15] M.-F. Moens, E. Boiy, R.M. Palau, and C. Reed. Automatic Detection of Arguments in Legal Texts. In *Proceedings of the International Conference on AI and Law (ICAIL-2007)*, pages 225–230, 2007.

[16] Marie-Francine Moens. Argumentation Mining: Where are we now, where do we want to be and how do we get there? In *FIRE '13 Proceedings of the 5th 2013 Forum on Information Retrieval Evaluation*, 2013.

[17] N. Naderi and G. Hirst. Argumentation Mining in Parliamentary Discourse. *Paper presented at 15th Workshop on CMNA, Bertinoro, Italy*, 2015.

[18] O. Onyimadu, K. Nakata, Y. Wang, T. Wilson, and K. Liu. Entity-Based Semantic Search on Conversational Transcripts Semantic. In *Takeda, H., Qu, Y., Mizoguchi, R. and Kitamura, Y. Semantic Technology. Lecture Notes in Computer Science (7774)*, pages 344–349. Springer, 2013.

[19] O. Onyimadu, K. Nakata, T. Wilson, D. Macken, and K. Liu. Towards Sentiment Analysis on Parliamentary Debates in Hansard. In *Kim, W., Ding, Y. and Kim, H.-G. Semantic Technology. Lecture Notes in Computer Science (8388)*, pages 48–50. Springer, 2014.

[20] Fabio Paglieri and Cristiano Castelfranchi. Trust, relevance, and arguments. *Argument & Computation*, 5(2-3):216–236, 2014.

[21] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up? Sentiment classification using machine learning techniques. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 79–86, 2002.

[22] Simon Parsons, Katie Atkinson, Zimi Li, Peter McBurney, Elizabeth Sklar, Munindar Singh, Karen Haigh, Karl Levitt, and Jeff Rowe. Argument schemes for reasoning about trust. *Argument & Computation*, 5(2-3):160–190, 2014.

[23] Andreas Peldszus and Manfred Stede. From argument diagrams to argumentation mining in texts: a survey. *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, 7(1):1–31, 2013.

[24] I. Rahwan, F. Zablith, and C. Reed. Laying the Foundations for a World Wide Argument Web. *Artificial Intelligence*, 171:897–921, 2007.

[25] Zaher Salah. *Machine learning and sentiment analysis approaches for the analysis of Parliamentary debates*. PhD thesis, University of Liverpool, 2014.

[26] Jodi Schneider. An Informatics Perspective on Argumentation Mining. In *ArgNLP*, 2014.

[27] Adam Taylor. Before Thatcher Came To Power, The UK Was Literally Covered In Gigantic Piles Of Garbage, 2013. `http://www.businessinsider.com/thatcher-and-the-winter-of-discontent-\\2013-4?IR=T` [Last Accessed: 02/02/16].

[28] Kristina Toutanova, Dan Klein, Christopher Manning, and Yoram Singer. Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network. *Proceedings of HLT-NAACL*, pages 252–259, 2003.

[29] Douglas N. Walton. *Fundamentals of Critical Argumentation.* Cambridge University Press, 2006.

[30] Adam Wyner, Raquel Mochales-Palau, Marie-Francine Moens, and David Milward. Approaches to text mining arguments from legal cases. In *Semantic processing of legal texts*, pages 60–79. Springer Berlin Heidelberg, 2010.