LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

DISSERTATIONEN DER LMU

39

KATSIARYNA STALPOUSKAYA

Automatic Extraction of Agendas for Action from News Coverage of Violent Conflict

Katsiaryna Stalpouskaya

Automatic Extraction of Agendas for Action from News Coverage of Violent Conflict

Dissertationen der LMU München

Band 39

Automatic Extraction of Agendas for Action from News Coverage of Violent Conflict

von

Katsiaryna Stalpouskaya

Universitätsbibliothek
Ludwig-Maximilians-Universität München

Herausgegeben von der
Universitätsbibliothek der Ludwig-Maximilians-Universität
Geschwister-Scholl-Platz 1
80539 München

Mit **Open Publishing LMU** unterstützt die Universitätsbibliothek der Ludwig-Maximilians-Universität München alle Wissenschaftlerinnen und Wissenschaftler der LMU dabei, ihre Forschungsergebnisse parallel gedruckt und digital zu veröffentlichen.

Erstveröffentlichung 2020
Zugleich Dissertation der LMU München 2019

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet abrufbar über http://dnb.dnb.de

Herstellung über:
readbox unipress
in der readbox publishing GmbH
Rheinische Str. 171
44147 Dortmund
http://unipress.readbox.net

Open-Access-Version dieser Publikation verfügbar unter:
http://nbn-resolving.de/urn:nbn:de:bvb:19-258073

978-3-95925-146-4 (Druckausgabe)
978-3-95925-147-1 (elektronische Version)

Contents

List of abbreviations

A4A	agenda for action
CAB-model	cognition-affect-behaviour model
CAF	collective action framing
CWC	Chemical Weapons Convention
DT	decision tree
DTR	distributional term representation
EU	European Union
FN	false negatives
FP	false positives
HC	the House of Commons
IDF	inverse document frequency
ISIL	Islamic State of Iraq and the Levant
*k*NN	k-nearest neighbours
LDA	latent Dirichlet allocation
LF	linguistic features
ML	machine learning
MLP	multilayer perceptron
NATO	North Atlantic Treaty Organization
NB	naïve Bayes
NGO	non-governmental organization
NLP	natural language processing
NYT	New York Times
OPCW	Organization for the Prohibition of Chemical Weapons
POS	part of speech
SME	small and medium enterprises
SVM	support vector machine
TF	term frequency
TN	true negatives
TP	true positives
TPB	theory of planned behaviour
TRA	theory of reasoned action
UK	United Kingdom
UNDP	United Nations Development Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization

UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations International Children's Emergency Fund
US	United States

List of tables

List of figures

Acknowledgement

Looking back at the time when I was working on this thesis, I realise that it has been a roller-coaster of events and emotions: I recall being appreciated and criticised, setting up deadlines and not meeting them, achieving results and celebrating successes, failing and learning from mistakes, maturing as a researcher and as a person, being desperate and determined, sharing my findings with peers at conferences, being supported and encouraged by some and challenged by others… I would not be able to accomplish this project without many people, who have been with me along the way.

I would like to express my sincere gratitude to everyone who has supported me on this scientific journey. First of all, these are my supervisors – Prof. Dr. Romy Fröhlich and Dr. Christian Baden. They have always been there for me with their wise advice, patience, timely feedback, constructive criticism and full support. They taught me to think critically, to be open to new ideas and not to be afraid. I can heartily call them my 'doctoral parents', without whom I would not have succeeded as a researcher.

Secondly, I extend my heartfelt thanks to all members of the INFOCORE research consortium. The core concept of this thesis – 'agenda for action' – originated within FP7 EU-INFOCORE project. I am grateful for the help constructing theoretical framework as well as collecting data used in this work. I would like to mention the team leads of content analysis working group – Prof. Dr. Romy Fröhlich, Prof. Dr. Kerren Tenenboim-Weinblatt, Prof. Dr. Rosa Berganza and Dr. Dimitra Dimitrakopoulou – and all their team members.

Thirdly, I would like to warmly thank Ludwig-Maximilians-Universität München, in particular Institute for Communication Science and Media Research (IfKW), for creating and providing friendly and inspiring research environment and letting me be part of it. I would like to extend a special appreciation to Dr. Manuel Wendelin, without whom the defence of this dissertation would have been impossible.

Fourthly, special shout-outs go to Tom Sharkie, Dr. Marc Jungblut, Dr. Yuan Zeng and Dr. Clarissa Schöller for proofreading, correcting and translating my thesis.

Last but not least, I would like to thank my family and friends who have been patient and helpful during these years.

for my parents

Zusammenfassung der Arbeit

Mediatisierte Kommunikation durchdringt viele Lebensbereiche und nimmt dabei vielfältige Formen an. Soziale Medien, strategische Kommunikation wie auch traditionelle Massenmedien informieren das Publikum nicht nur über den Zustand der Welt, sondern erfüllen auch eine Referenzfunktion, indem sie relevante Themen auswählen sowie einzelne Themen(aspekte) einordnen und betonen. Die mediale Selektion und Darstellung von Themen beeinflussen dabei die Interpretation, Bewertungen und Einstellungen der Menschen zu den dargestellten Themen und Akteuren (Lee, 2010). Aber ist die Rolle der Medien darauf beschränkt Aufmerksamkeit zu lenken und Interpretationen zu beeinflussen? Diese Dissertation untersucht, ob es möglich ist, zukünftige Ereignisse und die Mobilisierung für kollektive Handlungen durch die Analyse der medial vermittelten Kommunikation vorherzusagen.

Um dies zu tun analysiert die Arbeit die Rolle von 'agenda for action' in der mediatisierten Kommunikation. Darunter versteht man semantische Informationen, die einen Aufruf zu gewünschten oder notwendigen Handlungen oder Entwicklungen beschreiben oder die charakterisieren, welche Handlungen oder Entwicklungen ungewünscht sind. 'Agenda for action' sind zudem zielgerichtet, da sich auf einen erwünschten zukünftigen Zustand beziehen. Semantisch ist eine 'agenda for action' dabei ein Vorschlag bzw. eine Aufforderung, die aus drei Bestandteilen zusammengesetzt ist. Erstens besteht eine 'agenda for action' aus einem Referenzpunkt, auf den die geforderte Handlung sich bezieht. Dieser kann sowohl ein Problem sein, das behoben werden soll oder auch ein wünschenswerter Zustand, der beibehalten werden soll. Der zweite Bestandteil einer 'agenda for action' ist der Handlungsaufruf. Drittens werden die geforderten oder zu unterlassenen Handlungen näher definiert. Alle drei Komponenten können in einer 'agenda for action' sowohl explizit erwähnt, als auch implizit angedeutet werden.

Das hier verwendete Begriffsverständnis von 'agenda for action' basiert auf den theoretischen Überlegungen des FP7 EU-Projekts INFOCORE[1]. Hier werden 'agenda for action' als „prospective discursive con-

1 www.infocore.eu, Grant Agreement No. 613308

structions that postulate specific goals which must still be achieved“ (Tenenboim-Weinblatt, 2013, S. 92) definiert.

Konzeptionell können ‘agenda for action’ dabei an der Schnittstelle zwischen drei Forschungstraditionen verortet werden, nämlich den kommunikationswissenschaftlichen Ansätzen Agenda-Setting und Framing, sowie der Sprechakttheorie innerhalb der linguistischen Forschungstradition der Pragmatik. Alle drei Ansätze vereint, dass sie versuchen ‘agenda for action’ in ihre theoretischen Überlegungen aufzunehmen, jedoch fehlt allen drei Traditionen gleichzeitig ein systematischer Zugang und eine passende Operationalisierung.

In seinem ursprünglichen Verständnis beschreibt Agenda-Setting das Übertragen von Themen-Salienz von den Medien auf das Publikum (Kiousis & McCombs, 2004; McCombs and Shaw, 1972; Takeshita, 1997). Dieses erste Level des Agenda-Settings untersucht also, ob die Medien beeinflussen können *über was* die Rezipienten nachdenken. Diese Perspektive wurde durch das zweite Level von Agenda-Setting ergänzt. Hier wird untersucht, ob die Medien beeinflussen *wie* die Rezipienten über Themen denken (Golan & Wanta, 2001; McCombs, Llamas, Lopez-Escobar, & Rey, 1998; McCombs & Shaw, 1993). Schließlich weisen einige Forscher (z.B. Becker, 1977; Ghorpade, 1986; Moon, 2014) auf eine Verbindung zwischen Agenda Setting und dem tatsächlichen Verhalten von Menschen hin. Auf dieser Verbindung baut die Idee der ‘agenda for action’ auf, in dem sie untersucht, ob die Medien beeinflussen können, wie die Menschen *handeln*.

Die zweite Forschungstradition, auf der das Konzept ‘agenda for action’ aufbaut ist Framing. In seiner viel zitierten Definition von Frames beschreibt Robert Entman (1993) die Handlungsempfehlung als eines der zentralen Bestandteile eines Frames. Das Konzept der ‘agenda for action’ kann als Operationalisierung der Handlungsempfehlung angesehen werden. Zusätzlich lässt sich auch eine Verbindung zwischen ‘agenda for action’ und der soziologischen Betrachtung von Frames herstellen. Hier betrachtet das Konzept der ‘collective action frames’, wie Frames eine kollektive Mobilisierung bewirken können, also wie Frames Menschen davon überzeugen etwas zu tun. Forschung zeigt dabei, dass Frames für soziale Bewegungen unterschiedliche Funktionen erfüllen können und so potentielle Anhänger für Themen mobilisieren. Die erste Funktion ist die 'Diagnose', sie identifiziert das zu behandelnde Thema. Die zweite Funktion ist 'Motivation', sie besteht aus zwei Teilfunktionen: zum einen der Aufforderung, sich an der gemeinsamen Aktion zu beteiligen und zum anderen der Identifikation einer Begründung für die Notwendigkeit der kollektiven Handlung. Schließlich soll die 'Prognose' mögliche Lösungswege aufzeigen. Laut Gamson (1992, 1995) werden diese Aufgaben innerhalb des Frames durch drei Komponenten erfüllt. Die Frame-Komponente 'Ungerechtigkeit' befasst sich mit dem menschlichen Bewusstsein über das Leiden und den Schaden, der durch ein Problem verursacht wird. Diese Komponente erfüllt die Aufgabe der Diagnose. Die zweite Frame-Komponente 'Agency' appelliert an das menschliche Bewusstsein und die Möglichkeit, die Situation durch kollektive Aktionen zu verändern. Sie ist für Teile der motivationalen Frame-Aufgabe verantwortlich, nämlich für ihre zweite Teilaufgabe, die die Grundlage für das gemeinsame Handeln bildet. Die dritte Frame-Komponente ist 'Identität'. Ihr Ziel ist es die Anhänger zu einer gemeinsamen Ingroup ('wir') gegenüber einer Outgroup ('die') zu vereinen. Wie jedoch mehrfach hervorgehoben wurde (Benford & Snow, 2000; Gamson, 1995; Gerhards & Rucht, 1992; Giugni, 2006; Sanfilippo et al., 2008; Snow & Benford, 1988, 1992), gibt es Diskrepanzen zwischen Frame-Aufgaben und Komponenten. Es fehlt eine klare Unterscheidung zwischen Prognose und Motivation, es gibt keine eigene Frame Komponente, die die Aufgabe der 'Prognose' erfüllt, und die Aufgabe des motivationalen Framing wird von der Frame-Komponente 'Agency' nur teilweise erfüllt. Daher können Framing-Aufgaben nicht eins zu eins auf Frame-Komponenten abgebildet werden. Die Einbeziehung des Konzepts der 'agenda for action' kann dieses Problem lösen. Diagnose und ein Teil der Motivation können in der 'agenda for action' zusammengeführt und der Handlungsfähigkeit zugeordnet werden. Motivation kann dann die Aufgabe erfüllen, eine Handlungsgrundlage zu liefern und auf Ungerechtigkeit abgebildet werden.

Die dritte Forschungstradition, auf der das Konzept der 'agenda for action' aufbaut, ist die linguistische Sprechakttheorie. Diese beschäftigt sich mit der Beziehung zwischen Sprache und Handlungen. Sprechakte werden dabei als Sätze verstanden, die eine Handlung ausführen, nur weil sie ausgesprochen werden. Austin (1962) und Searle (1975) trugen zu der Entwicklung einer Reihe von Kriterien und formalen Markern bei, die dabei helfen Sprechakte zu identifizieren. Eines der Hauptkriterien ist die Richtung der Anpassung. Hier unterscheidet die Forschung zwischen zwei unterschiedlichen Arten: von Wort zu Welt und von Welt zu Wort. Ersteres liegt vor, wenn das Ziel eines Sprechers darin besteht, die Welt gemäß dem Vorschlag zu verändern, z.B. '*Bitte, öffnen Sie die Tür*'. Von einer Anpassung von Welt zu Wort wird gesprochen, wenn ein Redner die Äußerung dazu bringen will, die Welt wirklich widerzuspiegeln, wie z.B. '*Ich bin so müde*' (vorausgesetzt, ein Redner ist tatsächlich müde). Während Searle (1975a, 1976) hauptsächlich die Fälle von Sprechhandlungen untersucht, die durch Sprechaktverben eingeführt werden, betont spätere Forschung (z.B. Bach, 1994, 2014; Holdcroft, 1994), dass Sprechakte auch durch Grammatikmarker wie Modalverben oder Imperativsätze ausgedrückt werden können. Sprechakte können auch implizit sein und erfordern einen breiteren Kontext, um illokutionäre Akte abzuleiten. Die Sprechakttheorie wird in der vorliegenden Arbeit mit dem Konzept der 'agenda for action' verbunden. So werden alle Möglichkeiten in der natürlichen Sprache berücksichtigt, die illokutionäre Akte mit der Anpassungsrichtung von Welt zu Wort auszudrücken. Zusätzlich beinhaltet die Analyse von 'agenda for action' alle Möglichkeiten, in denen in natürlicher Sprache zu Handlungen aufgerufen werden kann: direkt und indirekt ebenso wie wörtlich und bildlich. Dadurch können neue Erkenntnisse über Handlungsanweisungen gewonnen werden und bestehende Forschungstraditionen systematisierend verknüpft werden.

Für die Analyse von 'agenda for action' in der Textanalyse ist es zentral zu betrachten, welche Maßnahmen gefordert werden. Es gibt eine fast unendliche Anzahl von Möglichkeiten, 'agenda for action' zu klassifizieren. Welche Taxonomie verwendet werden soll, hängen vom Zweck der Analysen und der Aufgabendomäne ab. In der vorliegenden Arbeit extrahiere und analysiere ich 'agenda for action' aus der Medienberichterstattung über gewalttätige Konflikte. Deshalb beinhaltet die entwickelte Taxonomie u.a. Forderungen nach Eskalation, Deeskalation, Bestrafung und Unterstützung. Dieses Thema ist besonders relevant für die Analyse von 'agenda for action', weil Kriege und Konflikte

Ereignisse mit besonders hohem Nachrichtenwert sind und geforderte Handlungen hier häufig sehr folgenreich sind.

Eine Analyse der geforderten 'agenda for action' ermöglicht es dabei zukünftige Ereignisse vorherzusagen. Um Vertrauen in diese Vorhersagen zu gewinnen, müssen agendas for action aus großen heterogenen Textkorpora extrahiert und im vorliegenden Kontext interpretiert werden. Daher wird in der vorliegenden Arbeit eine automatisierte Inhaltsanalyse von großen Datenmengen durchgeführt. Die Studie verwendet dabei maschinelles Lernen (ML) um die Daten zu analysieren. ML ist definiert als das „field of study that gives computers the ability to learn without being explicitly programmed" (Samuel, 1959, S. 210). Um ein ML-basiertes Informationsextraktionstool zu implementieren, müssen drei wesentliche Schritte abgeschlossen werden. Zunächst müssen Daten für einen Algorithmus bereitgestellt werden, aus denen der Algorithmus wiederkehrende Muster erkennen und somit die Klassifikationsaufgabe 'erlernen' kann. Zweitens müssen die Daten in ein maschinenlesbares Format übertragen werden. Schließlich muss ein Lernalgorithmus entwickelt werden, der auf Grundlage der bereitgestellten Daten Schlussfolgerungen zieht und inferenzstatistische Analysen erstellt.

Erster Schritt

In der automatisierten Textanalyse werden die gesammelten Daten als Korpus bezeichnet. Für viele Aufgaben, einschließlich der automatisierten Extraktion und Klassifizierung von 'agenda for action' muss zuerst ein annotierter Korpus erstellt werden. Das bedeutet, dass jedem Datenelement in eine Korpus ein Label zugewiesen werden muss. Dabei wird markiert, ob ein Datenelement eine 'agenda for action' ist oder nicht, und falls es sich um eine 'agenda for action' handelt, wird festgelegt um welche Klasse bzw. welchen Typus von 'agenda for action' es sich handelt (z.B. Aufruf zur Eskalation). Der annotierte Korpus dient drei Zwecken: einen ML-Algorithmus zu trainieren, ihn auszuwerten und zu testen wie präzise ein Algorithmus die implementierte Aufgabe durchführt. Daher wird der Korpus in der Regel in drei Teilmengen unterteilt. Da es keinen vordefinierten Korpus von 'agenda for action' gibt, wird zuerst ein Korpus erstellt ('Kaggle', 2015). Dieser dient für die Umsetzung des ML-basierten Ansatz und um 'agenda for action' extrahieren und klassifizieren zu können. Um das gesamte Spektrum der medial vermittelten Kommunikation abzudecken, besteht dieser Corpus aus Texten von traditionellen Massenmedien, sozialen Medien, sowie aus strategischer und politischer Kommunikation. Das betrachtete Textmaterial stammt aus dem Datensatz des FP7 EU-Projekt INFOCORE, dessen Ziel es war, die Rolle der Medien bei gewalttätigen Konflikten zu untersuchen. Demzufolge beschreiben die Texte im Korpus ausschließlich Kriege und gewalttätige Konflikte. Die zu analysierenden Texte wurden zunächst in Sätze aufgeteilt und anschließend wurde jedem Satz ein Label zugewiesen, das beschreibt, ob es sich um eine 'agenda for action' handelt. Falls dies zutrifft wurde zusätzlich festgehalten, welche Art Handlung in der 'agenda for action' gefordert wird.

Zweiter Schritt
Nachdem die Daten gesammelt wurden, müssen sie in ein maschinenlesbares Format übertragen werden. Dieses Verfahren wird für gewöhnlich als die Extraktion von Features bezeichnet. Diese Features werden anschließend an einen Algorithmus weitergegeben, der darin Muster in den Daten erkennt und aus diesen Mustern Regelmäßigkeiten extrahiert, die er auf ähnliche Daten anwenden kann. Dieses Verfahren lässt sich als maschinelles Lernen bezeichnen. In der Computerlinguistik werden für gewöhnlich N-Gramme und das TF-IDF-Maß als Features extrahiert (Cavnar & John M. Trenkle, 1994; Fürnkranz, 1998). Bei N-Grammen mit N=1 oder sogenannten Unigrammen untersucht man jedes Wort als einzelnes Feature. Dies wird auch als 'bag-of-words' Ansatz beschrieben. Ist *N* größer 1, werden N benachbarte Worte als eine Entität angesehen und gemeinsam analysiert. Anschließend wird jedem N-Gramm ein TF-IDF-Wert zugewiesen. TF-IDF steht für 'term frequency – inverse document frequency'. Dieser Wert setzt das Vorkommen eines N-Gramms in einem Text im Verhältnis zu der Anzahl an Texten in denen dieses N-Gramm erscheint (Jones, 2004; Salton & Buckley, 1988; Salton & McGill, 1986; Wu, Luk, Wong, & Kwok, 2008). Zusätzlich zu den TF-IDF gewichteten N-Grammen, werden eine Reihe linguistischer Features betrachtet. Sie beschreiben grammatikalische und lexikalische Muster, die helfen 'agenda for action' zu unterscheiden. Diese Features überprüfen ob ein Satz lexikalische Marker wie Sprach-

aktverben, Modalverben oder spezifische Wortkombinationen enthält, wie z.B. '*die Zeit des Handelns ist gekommen*' oder '*es ist unvermeidbar*'. Zudem suchen die Features nach grammatikalische Indikatoren für 'agenda for action', wie zum Beispiel Imperativen.

Dritter Schritt

Im dritten Schritt wird ein Algorithmus entwickelt, der es ermöglicht 'agenda for action' zu klassifizieren. Insgesamt gibt es eine Vielzahl an möglichen ML-Algorithmen. Die Wahl eines passenden Algorithmus hängt von der genauen Aufgabe und der verfügbaren Datenmenge ab (Abu-Mostafa, Magdon-Ismail, & Lin, 2012; Conway & White, 2012; Hastie, Friedman, & Tibshirani, 2001; Segaran, 2007). Für die vorliegende Dissertation wurden fünf Algorithmen gegeneinander getestet: support vector machine (SVM), *k*-nearest neighbours, naïve Bayes, decision tree und multilayer perceptron. Scikit-learn Implementation des Algorithmus wurde dabei angewandt (Pedregosa et al., 2011).

Jeder dieser Algorithmen wurde mehrmals ausgeführt. Dabei wurde die Menge an Trainingsdaten und die berücksichtigten Features (z.B. mit nur TF-IDF weighted *N*-Gramme oder mit TF-IDF Werten und zusätzlichen linguistischen Features) variiert. Die Extraktion von 'agenda for action' geschah dabei in zwei Schritten. Zunächst wurde entschieden, ob ein Satz eine 'agenda for action' darstellt oder nicht. Anschließend wurden die Sätze, die als 'agenda for action' erkannt wurden klassifiziert basierend auf der entwickelten Taxonomie. In beiden Schritten zeigte der SVM Algorithmus die besten Ergebnisse.

Um zu zeigen, wie der entwickelte Algorithmus zur Extraktion und Klassifizierung von 'agenda for action' in Textanalysen verwendet werden kann, wurde das vorgestellte Verfahren auf die Medienberichterstattung zu der Chemiewaffenkrise in Syrien 2013 angewendet. Hierbei wurden Texte aus der *New York Times* (*NYT*) und dem *Guardian* analysiert. Die Situation in Syrien war zu diesem Zeitpunkt sehr unklar und widersprüchlich. Mehrere, häufig umstrittene und heftig diskutierte 'agenda for action' wurden von verschiedenen Parteien vorgeschlagen, die jeweils versuchten, die Ereignisse zu verstehen und eine wünschenswerte zukünftige Entwicklung in Syrien aufzuzeigen (Baden & Stalpouskaya, 2015b). Der Fall kann somit das volle Potenzial der automatisierten

Analyse von 'agenda for action' demonstrieren, da im Diskurs eine Vielzahl an unterschiedlichen 'agenda for action' vorliegen sollten.

Im Jahr 2013 können mehrere Schlüsselereignisse identifiziert werden. Der erste Chemiewaffenangriff fand am 19. März in Aleppo statt. In der Zeit bis August gab es viele Diskussionen und Mutmaßungen darüber, ob der Angriff tatsächlich stattgefunden hat, wer dafür verantwortlich gemacht werden sollte und ob eine militärische Intervention der richtige Weg zur Lösung der Krise ist. Am 21. August fand der zweite Angriff in Ghouta statt. Am 29. August stimmte das Unterhaus im Vereinigten Königreich gegen eine militärische Intervention zur Lösung der Krise in Syrien. Am 14. September trat Syrien der UN Chemiewaffenkonvention bei und erklärte sich bereit, seine Chemiewaffenbestände aufzugeben. Ende September veröffentlichten die Vereinten Nationen einen Bericht, in dem sie feststellten, dass Nervengift in Syrien verwendet wurde.

Eine erste mögliche Anwendung des vorgestellten Algorithmus ist die Betrachtung der Anzahl der 'agenda for action' im Zeitverlauf. Die durchgeführte Analyse zeigt, dass während Eskalationen die Anzahl der 'agenda for action' in der Berichterstattung zunimmt. In Zeiten ohne nennenswerte Ereignisse lässt sich eine Abnahme an 'agenda for action' in der Berichterstattung aufzeigen. Eine solche Analyse kann verwenden werden um die Situation vor Ort in Echtzeit zu überwachen und zukünftige Entwicklungen vorherzusagen.

Ebenso ermöglicht die Anzahl der identifizierten 'agenda for action' in der Berichterstattung eine Einschätzung der Intensität der politischen Debatte. Die Erhöhung der Gesamtzahl der 'agenda for action' in den Medien kann Ausdruck eines heftigen politischen Diskurses sein. Somit ermöglicht eine Analyse der 'agenda for action' im Zeitverlauf relevante Momente im politischen Diskurs zu identifizieren, um diese in detaillierteren (qualitativen) Analysen zu betrachten.

Durch die Analyse und den Vergleich der Veränderungen in der Qualität und Quantität der 'agenda for action' in der Berichterstattung ist es möglich, die Außenpolitik verschiedener Länder zu vergleichen, sogar in Echtzeit. Insgesamt zeigt die Analyse, dass es im Untersuchungszeitraum nur wenig Unterschiede zwischen dem *Guardian* und der *NYT* hinsichtlich der Anzahl der ausgedrückten 'agenda for action' in der Medienberichterstattung und ihrer Algorithmus-basierten inhaltlich Klassifizierung gibt. Dies suggeriert, dass der politische Diskurs in beiden Ländern eine ähnliche Außenpolitik fordert. Es zeigt sich, dass eine automatisierte Analyse von 'agenda for action' hilfreich sein kann für eine Betrachtung von politischem Wandel und Entwicklungen im Zeitverlauf. Der Algorithmus, der im Rahmen dieser Dissertation entwickelt wurde, ermöglicht es politische Entwicklungen in Echtzeit zu verfolgen, was besonders nützlich sein kann, wenn es keine Möglichkeit für tiefergehende qualitative Untersuchungen gibt.

Die vorliegende Dissertation eröffnet neue Möglichkeiten für die zukünftige Forschung. Die entwickelte Methode kann auf verschiedene Textkorpora und in verschiedenen Bereichen angewendet werden, z.B. können Empfehlungen zur medizinischen Behandlung aus medizinischen Texten extrahiert werden; To-Do-Listen können durch die automatisierte Extrahierung von 'agenda for action' aus einem E-Mail-Korpus generiert werden. Allerdings muss eine entsprechende Taxonomie entwickelt und der Algorithmus umgeschult werden.

Das Verfahren und der Algorithmus können ebenfalls verbessert werden. Der Trainingsdatensatz kann vergrößert werden, die Einbeziehung anderer linguistischer Merkmale oder die Verfeinerung der in der vorliegenden Arbeit verwendeten Lexika verbessert auch die Klassifikationsergebnisse. Es lohnt sich, andere Algorithmen auszuprobieren, z.B. Deep Learning und die Verwendung von Wortvektoren (Mikolov et al., 2013) als Feature. Trotz dieser Mängel erreicht diese Dissertation zwei große Ziele. Erstens demonstriert die Arbeit, dass es möglich ist, komplexe semantische Konstrukte wie 'agenda for action' automatisiert zu erfassen und zu klassifizieren. Eine solche automatisierte Analyse von komplexen Sprachmustern eröffnet eine Reihe von Möglichkeiten für die Zukunft von automatisierten Textanalysen in verschiedenen Anwendungssituationen und Forschungsdisziplinen. Zweitens zeigt die Dissertation wie ein interdisziplinärer Ansatz es ermöglicht, ein Objekt aus verschiedenen Blickwinkeln zu erforschen, neue Erkenntnisse zu generieren und alle angewandten Disziplinen voranzutreiben.

Introduction

According to Jakobson (1960), one of the main language functions is a conative function, which is directed towards the addressee of a message and is mainly realized via imperative statements. Indeed, the only way to have someone perform an action (or an inaction) is to communicate it to the intended actor. The act of communication can be performed in many ways: It can be uttered directly and literally, or put figuratively and indirectly, it can be spoken or written, it can even be transmitted through non-verbal body language and gesture. The key is that the imperative must be communicated. When the addresser or the addressee (or both) of a message is group of people, one speaks of a collective directive statement (Meijers, 2007). The area where collective directives are the main means of communication is the political sphere, where deciders and policy-makers address large communities and prescribe specific courses of action as policies. Not always have collective directives been a form of a policy, frequently they emerge as calls to protest, to boycott, to help. One could say that agendas are created by directive statements. Typically, the addressees of these agendas – groups of people – become aware of these calls for action not directly from their sources, but rather via different media, most commonly via mass media. In other words, collective agendas are usually realised as mediated directives (Kampf, 2013).

Since the 1960s, the branch of communication science called *agenda setting* studies political and collective agendas as they are set via mass media. Despite ample interest and many published works, agenda setting theory has traditionally been centred on issues or topics, involving the means by which mass media shapes human thought. The emergence of the theoretical framework of *framing* in the 70s (Goffman, 1974) made a step forward in the study of issues – namely, how media shapes people's attitudes towards them. Agenda was only sidetracked within this framework by Entman (1993) who claimed that the specifics of dealing with an issue are one of the prime frame functions. Action has been the focus of the branch called collective action framing (Gamson, 1992), which investigates how frames can mobilise large groups of people to act together. The scholarship exhibits no concerted attention of agenda setting theory to action in itself.

Pragmatics, more specifically *speech act* theory, studies the properties of imperative language. This theoretical framework studies those propositions that can perform an action just by the virtue of their utterance such as "*Go away!*" or "*I command you to leave*". However, speech act theory covers a broader range of such expressions, not only directives. The theory focuses on the property of language to perform an action rather than of the phenomenon of language use to effect change in the world around, the underlying psychological and linguistic mechanisms, and its social function.

This doctoral dissertation addresses the limitations described above by introducing a novel term 'agenda for action'. This concept is explained and combined with the theories of agenda setting, framing and speech act. I will show the benefit of combining these theories within the concept of 'agenda for action'.

I also demonstrate how agenda for action can be used for analyses of text. Namely, I show that agenda for action can be used as a political barometer and signal of policy changes. Drawing example analyses of news coverage of violent conflict, I show that agendas for action expressed in news can reflect and even forecast how the situation will unfold. Finally, I enrich the toolkit of content analyses with the methods derived from computational linguistic approaches to text analysis. I apply machine learning to extract and classify agendas for action in an automatic fashion, thus enabling the assessment of a large text corpus.

This thesis has the following structure:

- In the first chapter, I provide the theoretical background of agenda for action. I highlight existing gaps in agenda setting, framing and speech acts, and demonstrate that the concept of 'agenda for action' can overcome these limitations, bridging the three fields. I explain that the concept is rooted in the theory of reasoned action and cultural memories.
- In the second chapter, I focus on machine learning and its application in computational linguistics for text analyses. I also provide an overview of both communication science and computational lin-

guistics approaches to analyse agenda-for-action-like entities both manually and computationally.
- In the third chapter, I apply machine learning for the purpose of extracting agendas for action. I describe the corpus compiled for this purpose, the algorithms used for classification, as well as criteria used to distil agendas for action. I shed light on the training of different statistical models, their evaluation and selection on the basis of best performance. I also provide an example of the agenda for action taxonomy with respect to war and violent conflict.
- In the fourth chapter, I showcase the algorithm developed in the third chapter, applying it to analyse news coverage with the example of the Syrian chemical weapon crisis of 2013. I extract and classify agendas for action from the American *New York Times* and the British *Guardian* and demonstrate the multiple interpretations of this information.
- I then conclude the dissertation with a review of the salient points from each chapter and suggestions for future research.

Each chapter contains a conclusion that sums the main points of interest.

This work – being an interdisciplinary project – targets a wide range of readers. The broad audience has been one of the biggest advantages of this thesis, but also a considerable challenge. The interdisciplinary nature of the work might also pose some difficulties for readers who are not specialists. I believe that all the information necessary to follow the argument has been provided either in the text or via references. I am convinced, despite all the difficulties associated with its interdisciplinary nature, the thesis benefits from the prism of multiple disciplines, allowing the concept of 'agenda for action' to be grasped and comprehended thoroughly.

1 Theoretical roots of agenda for action

The role of mediated communication – be it traditional mainstream media, such as newspapers or television broadcasts, social media or strategic communication – is manifold and touches upon many spheres of life. Not only does it inform the audience about the state of the world, but it also fulfils a referential function by priming and emphasising certain topics. In turn, media guides people's judgments, evaluations and attitudes towards the narrated content (Lee, 2010). But is the role of media limited in its capacity to direct attention and shape thought? Could it also serve to motivate action – even to the extent that those agendas expressed in the media can be used to predict future events? In other words, can one predict collective action through the analysis of mediated communication?

In this chapter I introduce a new concept, termed 'agenda for action', which can fulfil the role of a 'fortune teller', predicting collective action and social mobilisation. I discuss the conceptual roots of this phenomenon, demonstrating that mainstream communication science approaches touching on collective action, namely agenda setting and framing, have thus far lacked clear application of the concept in their theoretical frameworks. The concept of 'agenda for action' also enables a bridge between communication science and linguistics: Both frameworks study language use to force people to act, but from different angles.

1.1 Agenda for action: introduction into the concept

An agenda for action is a piece of semantic information that contains a request, desire or call for a particular outcome. It may also contain directives for desired or avoidable actions to achieve a certain state of affairs. These can be the calls to change the current unfavourable or undesirable situation as well as agendas to maintain the status quo. In natural language, agendas for action normally take the form of a proposition, which primarily consists of three parts. First, the entity towards which the action is directed (it may be a problem, or a positive) needs to be stated (1). Second, there should be a specific expression imploring action (2). Finally, the required motion/behaviour should be defined (3). All three components may be explicitly mentioned in each proposition, but they also may be implicit, requiring derivation from the context or cotext. Moreover, the call to act should be addressed to the future. It does not directly translate into grammatical future tense or any lexical indicators of future orientation, but rather addresses the logic that one cannot ask for changes in the past or try to amend it. It is only possible to change or keep the current situation in order to have a desirable future. It means that reported speech, for instance, can also communicate agendas for action: '*They said in the interview yesterday, that they kept on insisting on the peaceful solution to the crisis.*' In this example, there are two formal indicators of the past: the adverb 'yesterday' and the verb 'kept' in past tense. However, this is a merely grammatical peculiarity. The crisis (which corresponds to the first component of an agenda for action – the issue) is not resolved yet and the desired outcome (peaceful solution, the third component of an agenda for action) has not been reached yet. So in fact, the proposition is addressing the future. In its canonical version, this agenda for action could look like '*We insist on the peaceful solution to the crisis.*'

The term 'agenda for action' is relatively new in communication science. It has been introduced and developed within the theoretical framework of FP7 EU-Project INFOCORE[1], which examined the role of mass media in conflict areas (Baden & Stalpouskaya, 2015a; Baden & Tenenboim-Weinblatt, 2016; Stalpouskaya & Baden, 2015). The research group has defined agenda for action as "prospective discursive constructions that postulate specific goals which must still be achieved" (Tenenboim-Weinblatt, 2013, p.92). A complete agenda construction consists of three elements: a presentation of the present state or dynamic that cannot justifiably be left to itself; a future state that is desirable and attainable; and a set of more or less specific courses of action suitable to progress from the lamentable or precarious current to the desirable future state (Baden & Tenenboim-Weinblatt, 2016; Benford & Snow, 2000).

1 www.infocore.eu, Grant Agreement No. 613308

Its affiliation with the INFOCORE research project has determined the thematic focus of the present study: agenda for action taxonomy, the extraction of agendas for action and analysis with respect to media coverage of war and violent conflict. Hence, one of the aims of present dissertation is to be able to use agendas for action to track the dynamic of a conflict and, perhaps, to foresee its development.

In the coming sections, I will shed light on the concept of 'agenda for action' and explain how it advances communication science and linguistics through a synergy with agenda setting theory, framing and speech act theory. I will also bring in studies in psychology about reasoned action to demonstrate the linkage and causal relations between agendas for action advanced in media texts and collective action.

1.2 Conceptual roots

1.2.1 Agenda setting theory

Agenda setting is one of the most considered branches of communication science. In 2005, it boasted more than 400 publications, including journal publications, conference papers, theses, and monographs (McCombs, 2005). The rise of agenda setting theory was determined by the seminal study of McCombs and Shaw, concerning American presidential elections (1972). The publication showed empirically that media affects public opinion, i.e. it determines to a great extent what issues people think *about*. This work laid the foundation of public agenda setting. Thereafter, media agenda setting and political agenda setting emerged that examine the sources of mass media and governmental agendas, in contrast to public agenda setting whose dependent variable is the public (Dearing & Rogers, 1988; Lang & Lang, 1983; Reese, 1991; Rogers, Dearing, & Chang, 1991; Shoemaker, 1989). Public, media and policy agenda differ from one another with respect to its target audience, i.e. who is the receiver of an agenda – the public, politicians or mass media (Denham, 2010). The meaning of agenda in all three cases was an issue or a topic that dominates discourses and minds. Agenda as an action was not considered by these frameworks.

At the time when agenda setting theory was developed, communication science was facing several problems, which were supposed to be handled by the newly emerging theory. One of the limitations was the absence of connections between different components of the theory: "… decades of research into persuasive effects on attitudes and behaviours had left many scholars frustrated. Attitudes were not clearly connected to behaviour, and media were not clearly and consistently connected to either" (Kosicki, 1993, p. 103). Naturally, in order to address those problems and to bridge together attitudes and behaviours, agenda setting theory has incorporated aspects other than issues into its scope: A number of researches have shown that except for prompting the audience what to think *about*, or issue agenda setting (Takeshita, 1997), media also have some impact on people's judgments and evaluations. This phenomenon has become known as second level or attribute agenda setting (Golan & Wanta, 2001; McCombs, Llamas, Lopez-Escobar, & Rey, 1998; McCombs & Shaw, 1993). For agenda setting theory, the advent of the second level heralds the transition of the theory to a state that was no longer merely a study of issues, but rather a study of moving salience from one instance to another (Kiousis & McCombs, 2004). This development, on the one hand, has broadened the scope of agenda setting, on the other hand, it has drawn scholarly attention even further away from the core meaning of agenda – action.

While studying the causes and effects of agenda setting, most scholarly attention has been payed to the former – to the premises and psychological roots of agenda setting (such as the need for orientation), to the reasons and circumstance behind agenda setting, as well as the format of the agenda (Lee, 2015; McCombs & Stroud, 2014; Weaver, 1977). The effects and results of agenda setting have typically been explored through the prism of its cognitive effects such as salience, importance or accessibility of an issue (Iyengar, 1990; Iyengar & Kinder, 2010 (1987); Kim, Scheufele, & Shanahan, 2002; Nelson, Clawson, & Oxley, 1997; Scheufele, 1999; Scheufele, 2000).

The conclusion of the studies dealing with the outcomes of agenda setting has generally been that it is a powerful tool, capable of bending public opinion or attitude. This result is achieved due to the ability of media to move focus and shape emphasis, thereby priming attitudes and affecting the preferences and affective states of the audience (Ghanem, 1997; Golan & Wanta, 2001; Kiousis & McCombs, 2004; McCombs, Lopez-Escobar, & Llamas, 2000; McCombs & Shaw, 1993; Takeshita, 1997).

The result of media shaping public opinion and attitude is relatively unattended in the scholarship. As some sporadic works have demonstrated (e.g., Becker, 1977; Kepplinger & Roth, 1979; Moon, 2008), the consequence of ideas and thoughts being largely mental, is that people behave and act upon the issues with respect to the affective effects of agenda setting. As mentioned, however, work in the field of behavioural outcomes of agenda setting is relatively scant compared to the volume of available work on agenda setting.

The pioneering work in this regard belongs to Becker (1977) who focused on how the salience of an issue influenced political behaviour. He found out that the participation in political campaigns was higher among politicians who placed importance on the Vietnam War. Kepplinger and Roth (1979) addressed the field of consumer behaviour and examined that media coverage of oil crises prompted people to buy more petrol products. Sutherland and Galloway (1981), followed by Ghorpade (1986) demonstrated the way agenda setting affects public behaviour in the field of advertisement. The former have confirmed previous findings reporting a positive correlation between agendas set by media and public activities. Ghorpade elaborated this finding, introducing a two-stage model: The first stage refers to the transfer of salience from advertising to public mind, and the second stage indicates the transfer of salience from public mind to behavioural outcome. The model was empirically supported by Roberts (1992), who concluded that "the mass media may not only tell us what to think about but they influence what actions we take regarding those thoughts" (p. 878). The first stage corresponds to the shift of salience explained by classical agenda setting theory, while the second stage has, so far, only been studied crudely. Weaver (1991), Stroud and Kenski (2007), and Kiousis and McDevitt (2008) moved the studies back to the field of political behaviour, re-confirming the existence of dependency between public agenda and collective action.

The cognition-affect-behaviour model (CAB-model) has been developed within the study of the effects hierarchy of agenda setting. It suggests that agenda setting effects may be split into three groups, where the effects of the next group depend and expand upon the antecedent. Cognitive effects that address people's need for information serve as basis and necessary premise for affective outcomes that build upon cognition. These effects are the result of moving salience from an issue toward attitudes. The first two groups determine behavioural outcomes (Berelson, 1996; Lavidge & Steiner, 1961; McGuire, 1986; Severin & Tankard, 2001). In the 2000s, Moon dedicated a number of works to investigating the behavioural outcomes of agenda setting. She has employed the CAB-model to investigate how media affect political actions. She demonstrated that the second affective step of the CAB-model, operating via second-level agenda setting, triggers strong attitudes toward candidates, which, in turn, leads to various types of political participation and involvement (Moon, 2008, 2013, 2014).

Despite being a very powerful, long-established and developed framework, agenda setting still has some gaps and flaws. Even though agenda setting theory emerged as the study that was supposed to bring together issues, attitudes and behaviours, only the first two have become its core focus, receiving ample scholarly attention. Agenda setting theory has failed to draw a clear distinction between issues and actions (Hardy & Sevenans, 2015; Helfer & Wonneberger, 2015). Very little work has been done on the behavioural outcomes of agenda setting and its influence on collective action (Bimber, 2017; Kosicki, 1993; McCombs, 2004; McCombs & Estrada, 1997; Moon, 2008; Snyder, 2017). In contrast, the work that has been done enables agenda setting theory to easily incorporate the concept of 'agenda for action' developed in this work, which will advance and contribute greatly to the theory as a whole.

1.2.2 Framing

Framing is another approach to news analysis, which, according to Bryant and Miron (2004), became more popular than agenda setting in the 2000s. Goffman (1974) introduced the concept of framing to the field of communications, defining frames as schemata which people use to make sense of events, to interpret news discourses and guide collective action. Twenty years later, Entman gave a definition of framing, which has become the widely accepted standard: "To frame is to select some aspects of a perceived reality and make them more salient in a communicating text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and / or treatment recommendation" (Entman, 1993, p. 52). Through this definition the action guiding function is named as one of the key frame elements. The overlaps with attribute agenda setting definition, namely drawing attention to moral evaluations of a problem, have caused a huge scientific debate regarding the relationship between these two frameworks (Pan & Kosicki, 1997). The debate has questioned whether framing is an extension of agenda setting, its second level (Ghanem, 1997; McCombs, Lopez-Escobar, & Llamas, 2000; Takeshita, 1997), or a related but still independent field (Kosicki, 1993; Scheufele, 2000). A counter-argument has been advanced that agenda setting has rather come to a 'dead end' by introducing attribute agenda setting and will thus dissolve in the field of framing or traditional persuasion research (Takeshita, 2006).

As the answer to this question does not impact current research, I will only mention that the ongoing dispute confirms that the fields of framing and agenda setting are interconnected and related to one another – it is hardly possible to talk about one, without mentioning another (Takeshita, 2006; Weaver, 2007). More relevant to this study, however, is that framing as a scholarly tradition has, in fact, paid more attention to the behavioural outcomes of salience shifts, when compared to agenda setting theory. This finding renders framing a highly relevant to the content of this thesis.

As suggested by Goffman's (1974) and Entman's (1993) definitions of framing, treatment recommendation is one of the frame's prime components. Moreover, Entman (2003) has claimed that among the four frame elements, the problem definition and the treatment recommendation are the most important. In other words, one of the frame's basic functions is to "*suggest remedies* – offer and justify treatments for the problems and predict their likely effects" (Entman, 1993, p. 52). However, most of the scholarly work on framing has focused on the other frame elements – problem definition, causal interpretation and moral evaluation – paying far more attention to the manner in which problems are presented in the media, with a focus on trying to explain the roots of disagreements evoked by different frames (Chong & Druckman, 2007; Scheufele, 1999; Tewksbury & Scheufele, 2009). All the while the behavioural outcomes of the frame seem to be largely neglected (Gamson, 2005).

1.2.2.1 Collective action framing

Within the wide range of framing approaches, one that has consistently upheld framing's link to action and agendas is collective action framing. This type of framing focuses on action-oriented narrative constructs that appeal to people's beliefs, hopes and needs in order to inspire and make them act (Benford & Snow, 2000, p. 614; Gamson, 1995; Gerhards & Rucht, 1992; Giugni, 2006; Sanfilippo et al., 2008; Snow & Benford, 1988, 1992). Collective action framing has developed into a separate branch of the classical framing approach, whose core interest is the *action* stimulated by frames. According to this approach, the effect of action mobilization is achieved by fulfilling three 'core frame tasks' (Snow & Benford, 1988). The first task is diagnosis – identification of the problem to be acted upon as well as its roots. The second task is termed prognosis – the advancement of possible strategies to solve the issue. The final task is motivation – the call to engage in the collective action as well as to provide a rationale for doing so. For instance, *'the unions encouraged all the employees to join the national strike to get better working conditions'* – in this example the diagnosis is unsatisfactory working conditions, the prognosis is to strike, which will remedy the issue of poor working conditions if all workers participate, and the motivation is encapsulated in the word 'encouraged' with the rationale ameliorating work conditions.

However, defined in this way, there lacks a clear distinction between prognosis and motivation. The purpose of prognosis is to suggest specific courses of action (to strike in the example above), the aim of motivation is to express the need to act (embedded in the verb 'encourage' in the example above). Nonetheless, it is impossible to suggest a solution to a problem without calling for (or against) this solution. In other words, the task of prognosis cannot be fulfilled without accomplishing a motivational task and vice versa. The mere fact of proposing something implies either a call to pursue the suggested course of action or to avoid it. The opposite also holds; if there is a call for mobilization, there needs to be a specified course of action. This point is also reflected in language. It would be ungrammatical to say, 'we encourage' or 'they must', there needs to be a complement specifying what is needed or called for either in the same sentence, or inferable from the context. Thus, there exists an overlap between prognosis and motivation, though they may be merged into one task. Motivation also confounds the call to act with its rationale, these two aspects are rather independent from each other and hence should be split into different tasks: one corresponding to the reason only and another to the call to mobilise.

These discrepancies reflect inconsistencies between the aforementioned tasks and collective action frame components (Gamson, 1992, 1995). There exist three collective action frame components: Injustice – the part of the frame addressing human consciousness regarding suffering and harm. This component is traditionally associated with diagnostic framing, i.e. with the problem that needs to be acted upon (Anheier, Neidhardt, & Vortkamp, 1998; Klandermans & Weerd, 1999). The second frame component is agency – a frame part appealing to human consciousness and the possibility to change the situation through mobilisation and collective actions. This frame component is linked to the motivational framing task, namely to the part which provides the rationale for collective action rather than to the call itself. In the example above, poor working conditions would be the injustice component, while agency is the possibility of improvement. The third frame component, – identity – which is supposed to unite people as 'we' as opposed to 'they', clearly does not correspond to any framing task. This lack of congruence leaves prognostic framing without a dedicated frame component and the part of the motivational framing task that calls to action being only partially fulfilled by agency.

As a branch of framing, collective action framing has contributed the most to elucidating collective mobilization. It has also developed a theoretical framework to explain the frame structure and the mechanisms triggering collective actions. Yet, there exist discrepancies and gaps, namely a very vague border between prognosis and the part of motivation that calls to action. There are also inconsistencies between framing tasks and frame components. The concept 'agenda for action' introduced in this dissertation can address these issues and suggest possible solutions to them.

1.2.3 Agenda for action and speech acts

By definition, the second necessary component of agenda for action requires the presence of a specific expression of a need to perform an action. It addresses the linguistic aspect of agenda for action and touches upon several fields, i.e. grammar, semantics and pragmatics. From the viewpoint of meaning, agenda for action communicates a desire, hope, call or request to change or maintain the current situation. The peculiarity of such sentences is that the act of urging something is performed by virtue of its utterance. For example, '*Thank you for the present*' – the aim of the speaker is to express the gratitude for the gift she has received, the aim is being achieved solely by vocalising (or writing down) this sentence. The completion of the action stated in such sentences does not require any additional efforts or motions, in excess of making it known to the receiver. This phenomenon is studied by the field of linguistics called pragmatics, or more precisely, speech act theory (Sadock, 1974; Sbisà, 1995; Smith, 1990; Tsohatzidis, 1994). The theory suggests that virtually any utterance can contain an intention, i.e. embed an illocutionary force, and thus perform a perlocutionary act, i.e. make the audience fulfil an action or inaction (Austin, 1962; Cohen, 1973; Green, 1999). Therefore, in principle, any proposition may contain a call to action or a request for a change, expressed directly or indirectly, literally or nonliterally, explicitly or implicitly (Bach, 1994, 2014; Holdcroft, 1994; McGowan, Tam, & Hall, 2009; Searle, 1975b) (see also Bertolet (1994) for critics).

Even so, not every speech act implies an agenda for action. In other words, not each utterance that performs an action also calls for it. There exist numerous approaches to speech act classification. The class/classes comprising agendas for action depend greatly on the taxonomy used.

One of the pioneers in the speech act theory, John Austin (1962), suggested distinguishing five types of speech acts as presented in Table 1:

Class	Function	Examples of marker verb
verdictives	to perform a judgment, give an opinion or verdict, but not in a legislative sense	acquit, hold, calculate, analyse, estimate
exercitives	to communicate a decision in favour of, or against, a certain course of action	order, command, direct, plead, beg, recommend
commissives	to commit the speaker to a certain course of action	promise, vow, pledge, covenant, contract, guarantee, embrace
expositives	to fit arguments or opinions in the course of conversation	affirm, deny, emphasize, illustrate
behavitives	to express reactions and attitudes to events or other people's behaviour	apologize, thank, deplore, commiserate

Table 1: Speech act taxonomy by Austin

The weak points of Austin's taxonomy, as he admitted himself, are the lack of clear criteria for defining each class. It was created rather intuitively, without sound grammatical or semantic foundations. Also, the provided taxonomy classifies speech act verbs, rather than speech acts themselves. Speech act verb classification is language specific, i.e. it will be different for English, German, Russian, and so on. Speech act classes, conversely, are a universal characteristic of language (Langue as defined by Saussure (1916 (1959)), as such, they hold for all natural languages and are not bound by any one of its realizations.

Searle (1975a; 1976), addressing the abovementioned shortcomings, has amended Austin's taxonomy. He suggests assigning a different category to a speech act based on a number of criteria. The most important criterion, regarding the present study, is the direction of fit, which is the direction of illocutionary force of an utterance: making the world fit the word or vice versa (Anscombe, 1957; Humberstone, 1992). Searle has described five types of illocutionary acts presented in Table 2:

Type of illocutionary act	Direction of fit	Example	Correlation with Austin
representatives	From world to word	n/a	expositives, verdictives
directives	From word to world	invite, suggest, command	exercitives
commissives	From word to world	promise, vow	commissives
expressives	No direction of fit	thank, congratulate, deplore	behabitives
declarations	No direction of fit	n/a	n/a

Table 2: Illocutionary acts taxonomy by Searle

Two types of illocutionary acts in Searle's taxonomy do not have any marker verbs and illocutionary force is mainly expressed by the propositional meaning. These are representatives and declarations. Representatives usually include boasts and complains. For a proposition to perform an illocutionary act, context plays key role. The sentence '*We are the greatest nation*', to express a representative illocutionary act, has to be uttered under certain circumstances, with certain intonation and with an intention to show one's nation superior to others.

Another illocutionary act without marker verbs typical for the class are declarations. The necessary condition for them is that the utterer should be entitled to perform such acts, i.e. '*I appoint you a chairman*' – this speech act can be performed successfully only if a person with respective authority announces it. The same holds for '*I sentence you to 10 years' imprisonment*'.

Other taxonomies have been developed since Austin and Searle (e.g., Bach & Harnish, 1979; Sadock, 1994), but all of them built upon the taxonomies considered in this chapter, they complement each other, with many overlaps. Searle, however, has provided the soundest classification, in my opinion, backing it up with the definition of requisite criteria for introducing each class. Moreover, all of the speech verbs show different grammatical behaviour, i.e. they all have different surface syntactical representation (Harnish, 1994; König & Siemund, 2007; Portner, 2004; Sadock, 1994; Searle, 1976; Verschueren, 1980). Searle's taxonomy accounts for grammatical differences as well, namely, speech acts of one class, on the one hand, show similar grammatical behaviour and surface syntax, on the other, they differ in this regard from the speech act verbs of other classes. In light of this, I will draw upon Searle's classification as I distil agendas for action.

The key purpose and characteristic of an agenda for action is to *achieve a desirable future*. In other words, to make the future world fit one's desires, hopes, or expectations, one's words. Hence, the direction of fit as the criterion to classify speech acts can also be applied to agendas for action. Searle introduces two types of speech acts whose direction of fit is to make the world fit the word, these are directives and commissives. The two speech acts can be seen as two sides of one coin, as directives are committing the hearer to a specific course of action, while commissives commit the speaker to a specific course of action. Given Searle's definition, directive and commissive speech acts are considered to be agendas for action. Consider the two following sentences: '*We will not fight again, we **promise** you*' is an example of a commisive speech act and '*I **do not command** you to attack, I **command** you to die*' is a directive one. Both propositions are agendas for action.

There is one class in Searle's classification that does not possess any direction of fit, namely expressives. They convey a wide range of meanings, such as thanking, apologising, condemnation, condolence, welcoming, among others. Depending on the cotext and context, these propositions can serve as agendas for action, or at least can be understood as such. Given the focus of the present study on war and violent conflict, as well as the conventional property of illocutionary force, I have included some instances of expressives, for example blame and praise, in the scope of agendas for action. The intuition is that, if someone condemns the actions, words or intentions of others, she implies that it should not be done or said, e.g., '*We **condemn** such a motion*' equals to '*They **should not** act so*'. Thus, the illocutionary force of the utterance at hand equals the force of a proposition that clearly communicates the desire to change the current situation. Thus, it can be said that the former expression indeed possesses the direction of fit that aims to make the world fit the word, hence, it can be seen as an agenda for action. Sentences that welcome or praise follow the same logic. For example, '*I **encourage** you to vote for Dr. Michael Thompson on April 3rd*' can be rephrased as '*You **should** vote for Dr. Michael Thompson on April 3rd*' which clearly calls to vote for a specific candidate, with the desired outcome of Dr. Thompson's victory. In other words, such an exhortation attempts to make the world fit the word.

Speech acts expressing apology are also expressives. Unlike blame and praise, however, they can hardly communicate an agenda for action. In certain circumstances the utterance '*I am sorry for being late*' can be read as a commitment not to be late in the future, but in order to arrive to such a conclusion, the greater context is required, including intonation, body language and the whole conversation, as well as the extra-linguistic context. Moreover, apologies, being chiefly based on emotions, do not play an important role in understanding and eventually predicting conflict dynamics or tracking in real time. For this reason, apologies are not counted as agendas for action and have been excluded from the analysis. Only expressives with the meaning of condemnation and praise are considered as agendas for action, for the purposes of the present study.

1.3 Agenda for action: elaborated definition

In previous sections we have seen that several scholarly traditions in both communication science and linguistics have pointed out the mobilisation function of language, and that it is generally possible to 'do things with words' (Austin, 1962). This last property in particular is widely used in media texts. However, as far as concerns the theories of agenda setting and framing, action or behaviour as the outcome has been only briefly been mentioned, mainly vis-à-vis definitions (e.g., Entman's (1993) definition of framing). Action or behaviour itself has never become the focus of any of these theoretical approaches. Nevertheless, both theories can benefit greatly by incorporating the concept of 'agenda for action' within their scope.

Two models that take behavioural outcomes into account have emerged from within the agenda setting framework. The first was suggested by Ghorpade (1986) who has investigated agenda setting in the field of advertising. The model consists of two stages. In the first stage the salience moves from advertising to the mind of the audience, while in the second stage it affects people's actions. A more general, not domain-bounded CAB-model also accounts for behavioural outcomes of agenda setting (Berelson, 1996; Lavidge & Steiner, 1961; McGuire, 1986; Severin & Tankard, 2001). While the first stage of Ghorpade's model and the first two elements of the CAB-model have been thoroughly studied, the last component lacks theoretical background and operationalization. The present study aims to fill this gap, as the concept of 'agenda for action' provides a theoretical basis and operationalization strategy by which behavioural outcomes are incorporated into the scope of agenda setting.

As the theory mainly focused on salience, agenda setting theory has also investigated the movement of salience to issues, but also to attributes, thus initiating the branch of classical issue agenda setting (Takeshita, 1997) – attribute, or second level, agenda setting. Given this, the agenda for action can be seen as a part of agenda setting. It may be regarded as a special case, where agenda setting determines the issue, while agenda for action prescribes the relevant action. It can be also considered an extension of agenda setting wherein issue agenda setting defines the issue – what to think *about*, attribute agenda setting sets the attitude towards it – *how* to think about it, finally action agenda setting, or agenda for action, determines the course of action to deal with the issue – what to *do* about it[2].

If agenda setting theory is restricted to its classical version, to the theory that investigates the cognitive effects of agenda setting, and if one considers the affective and behavioural outcomes as belonging to framing, then the concept of 'agenda for action' can be seen as an operationalization of 'treatment recommendation' the fourth frame component introduced by Entman (1993). Agenda for action can also advance

2 The inclusion of behavioural outcomes into the scope of agenda setting model also makes sense from etymological perspective: Both terms 'agenda' and 'action' derive from the Latin *agere* – 'to do' (Onions (1996), it shows that action is a necessary part of any agenda and the study of agenda setting is incomplete without bringing action into its scope. This is also reflected in the meaning of the word 'agenda' in modern English. Oxford Dictionary of English (2010) defines 'agenda' as "a list of items to be discussed at a formal meeting: *The question of nuclear weapons had been removed from the agenda*", with the first subsense defined as "a plan of things to be done or problems to be addressed: *He vowed to put jobs at the top of his agenda*" (p.31). The same holds for a set expression 'set an agenda' which according to the same dictionary means "to influence or determine a program of action: *'He has set the agenda for future work in this field'*".

the approach of collective action framing by eliminating the inconsistencies between frame structure suggested by Gamson (1992; 1995) and collective action framing tasks (Snow & Benford, 1988). The outcome is that these approaches are linked. The part of the frame, though absent from Gamson's original description, but which could unite the action initiating tasks of the collective action frame – diagnosis and motivation – is the concept of 'agenda for action'. It calls to action, while precisely specifying the directive.

Including agenda for action in Gamson's frame structure enables it to explain *how* exactly collective action is mobilized. Moreover, in the context of social movements, it is of great importance to be able to identify any actions that may follow; whether a protest will be peaceful, whether workers will strike, or whether there will be violent armed riots. To study agenda for action in this regard can help to predict collective action, as well as to better deal with social movements, either facilitating negotiations, or meeting the demands of protesters, or preventing harm and negative consequences. Integrating the concept of 'agenda for action' into the collective action frame structure generates new avenues for research: It might shed some light on the relations and connections between frame components; such as defining whether people more eagerly stand up for violent protest when the 'identity' part is triggered opposing 'us' to 'them'.

The field of linguistics concerning 'how to do things with words' is a branch of pragmatics – speech act theory, which can also benefit from the work of the present dissertation. This thesis presents a systematic and exhaustive study of how calls to action can be expressed in natural language, going beyond directive and commissive speech acts (Searle, 1975a, 1976). It draws attention and includes into its scope the instances that traditionally have been studied in disparate scholarly traditions. The analyses accomplished in this dissertation bring together direct, literal illocutionary acts that can be performed by the means of 'force-indicating' devices or 'force-markers', such as sentential moods, as in the grammatical imperative '*Fight them*', performative verbs '*I command to fight them*' or modal verbs '*They must obey*', with implicit and indirect speech acts, therefore including expressives in its scope.

As one of the goals of the present research is to develop a computer program for the automatic extraction and classification of agendas for action from texts, it is essential to draw a clear border between the propositions that communicate an agenda for action and those that do not. Given the focus of my thesis on media coverage of war and violent conflict, it is important to capture all nuances in meaning, including those propositions that merely hint at an agenda for action. Therefore, a number of indirect, nonliteral and implicit speech acts whose direction of fit is from word to world have been considered as agendas for action. For instance, if the speaker expresses her dissatisfaction with the current situation or claims that it should not or cannot continue in such a way, she most certainly wants the situation to change and is either calling upon someone to do so or committing herself to some action, e.g., '*We cannot leave it like this*'. These types of speech can also be expressed as a rhetorical question, which subtly implores the hearer to do something in order to achieve the desirable result. For instance, '*can we accept this attitude?*' If defining agenda for action as a proposition suggesting a course of action in order to solve the issue, then it can be communicated as an utterance about the means to overcome the problem, without explicitly naming the motion, as in, '*We see the war as the only solution*'. Finally, the dissatisfaction can be communicated in general terms, without concrete suggestions or calls: '*Some actions should be taken against the terrorists*'.

With certain assumptions, there is a finite number of ways in which to express word to world fit in natural language. It is constructions made under such assumptions that I consider to be agendas for action. These constructions are as follows:

- directive, commissive and partially expressive speech acts as defined by Searle (1976) and Austin (1962) with an enhanced speech act verb list (Wierzbicka, 1987);
- expressions with indirect illocutionary force: *'I would like to continue the meeting, please';*
- imperative sentences: *'Fight them';*
- propositions with modal verbs obliging someone to do something: *'They must obey';*

- propositions expressing this dissatisfaction of the speaker: *'We will not stand this aggression';*
- general expressions that a situation must not stand: *'Something should be done';*
- rhetorical questions: *'Can we accept such a treatment?';*
- propositions about desirable states: *'Peace is the only answer'*.

In light of the theoretical data presented thus far, agenda for action can be defined in the following way. An agenda for action is a proposition that expresses a call, a hope, or a desire for a specific action (or inaction) to be performed in order to change the present unfavourable situation or maintain the status quo in order to obtain a desirable future. Agendas for action can be expressed explicitly, through lexical or grammatical indicators of illocutionary force, or implicitly when apparently dependent on the cotext and context. They consist of three elements: (1) a definition of the current state or problem to be acted upon, (2) a specific expression of the need to act, and (3) a desired course of action or inaction. Accordingly, agenda for action stand at the intersection of three scholarly traditions: agenda setting, framing and speech act theory.

Agenda for action stems conceptually from the three theoretical frameworks described above, at the same time they advance and contribute to each addressing some of the identified gaps and flaws. Agenda setting, framing and speech act theory can be used as meta-theories to describe the structure of agenda for action. The first element of an agenda for action refers to a state or issue to be acted upon, which is the task of classical issue agenda setting; the second component requires a specific expression of the call to act to be present – this is fulfilled by studies addressing explicit and implicit speech acts (Searle, 1975b); finally, the third component prescribes the manner in which a state or an issue is managed – to answer to such a question has traditionally fallen within the scope of framing.

1.4 Agenda for action within the frameworks of cultural memories and reasoned action

Cultural memory, namely collective prospective memory, and the theory of reasoned action are also concerned with people's actions and behaviours. In this section, I describe the tenets of each theory, as relevant to agenda for action.

1.4.1 Prospective memories

The idea of 'agenda for action' is strongly linked to, and derives from, the concept of prospective memories (Ellis, 1996; McDaniel & Einstein, 2007; Meacham & Leiman, 1982; Morris, 1992; Tenenboim-Weinblatt, 2013). McDaniel and Einstein have defined prospective memory as "remembering to carry out intended actions at an appropriate time in the future" (McDaniel & Einstein, 2007, p. 1). Prospective memory, or remembering what needs to be done, complemented by retrospective memory, which is the remembrance of what has happened, together comprise collective memory, the former part of which addresses the future and the latter, the past (Ellis & Cohen, 2008; Meacham & Leiman, 1982). The concept of collective prospective memories has been introduced as a potential link between the theoretical frameworks of the past oriented collective retrospective memory and the future addressed agenda setting. Collective prospective memory can be seen as a 'collective to-do list', which can include, for instance, tasks in the field of education reform or military intervention (Tenenboim-Weinblatt, 2013, p. 11).

This is where the idea of 'agenda for action' flourishes. While addressing the future and being closely related to agenda setting, it conveys how exactly public to-do lists are created, in other words, how public agendas are built and the tools that are used to prescribe specific courses of action. Agenda for action exceeds the study of collective prospective memories in that it might not only 'remind' societies of existing agendas, but can also set completely new agendas and call for specific treatment in their regard.

1.4.2 Reasoned action approach

An agenda for action, of course, does not only call for specific action to be taken, but can in fact motivate actual behavioural change. The reasoned action approach (Fishbein & Ajzen, 2010) is a further development in the study of human behaviour prediction and merges the theory of reasoned action (TRA) (Ajzen & Fishbein, 1980; Fishbein, 1967; Fishbein & Ajzen, 1975; Hale, Householder, & Greene, 2002) and the theory of planned behaviour (TPB) (Ajzen, 1985, 1991; Madden, Ellen, & Ajzen, 1992). The main idea of all three frameworks is that the behaviour is determined and can be predicted by behavioural intentions. No behaviour will be performed unless pre-existing intention is present. Behavioural intention, in turn, is a function of attitude toward the behaviour, normative beliefs, and subjective norms, according to TRA. The model does not deny the importance of other background factors, such as resources and capacity along with emotions, education, age, and gender, among others, but those are secondary to intention and its determinants. Being domain independent, the model is able to predict and to change behaviour in any sphere and has proved its power in such fields as political, organizational, health care and discriminative behaviour.

The reasoned action approach may be regarded as a link that bridges agenda for action and collective actions. It describes what exactly happens and the mechanisms that are triggered between the setting of the action agenda and actual performance of the action. As part of the agenda setting chain (issue agenda setting – attribute agenda setting – action agenda setting), agenda for action takes an active role in shaping and generating human norms, beliefs and attitudes, as has been shown in the agenda setting literature (Ghanem, 1997; McCombs & Evatt, 1995). Norms and beliefs established by agenda setting, in turn, determine the behavioural intentions of the audiences, which is the single predictor of the performed behaviour. Knowing the public agenda set by media and which actions are being called for via agenda for action enables us to foresee the dominant attitudes and beliefs in society. With this knowledge, it may become possible to predict the mobilization of consensus as well as collective action. Similarly, if one wants to control or change behaviours, it can be accomplished by setting specific agendas and promoting certain agendas for action, thus influencing and changing collective beliefs and controlling human behaviours.

1.5 Conclusion

The link between uttered speech and performed behaviour has been addressed by multiple theoretical frameworks. Both agenda setting and framing have determined that media can direct not only about the focus of thought but also influences people in their actions regarding these thoughts. Pragmatics has examined the ways language can have someone perform desired actions. Introducing the term of 'agenda for action' to these fields adds theoretical basis and substantiates the concept which previous sporadic works only mentioned briefly. The three frameworks can also be used on meta-theoretical level to explain the inner structure of agenda for action. Additionally, the concept of 'agenda for action' has been shown to be loosely related to the scholarship on collective prospective memories and the reasoned action approach. The relation of agenda for action to all these disciplines and their place within humanities and social sciences is reflected in Figure 1:

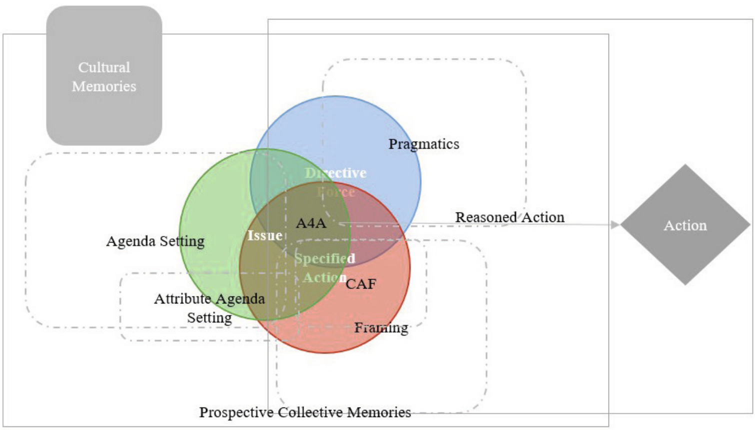


Figure 1: Agenda for action within humanities and social sciences
Where A4A stands for agenda for action, CAF – for collective action framing

As shown in Figure 1, the concept 'agenda for action' has emerged from the intersection of two mainstream approaches in communication science, namely agenda setting and framing, and the branch of pragmatics called speech act theory. The three theories are represented by three rectangles with dashed borders. Agenda for action can be seen as a part of the chain comprising issue agenda setting – attribute agenda setting – action agenda setting. This dependency is represented by the overlapping zone between agenda setting and agenda for action. The linkage between attribute agenda setting and agenda for action most likely exists, though requires further study and empirical investigation, which is not the part of the current research. For this reason, the diagram shows no overlap between the two respective nodes. Agenda setting, conversely, provides the first necessary component of an agenda for action, namely, the issue to act upon, which is shown by the green circle. Agenda for action together with attribute agenda setting are necessary premises for crowds to engage in collective action, as has been shown by the reasoned action approach. This connection is represented by an arrow from agenda for action to action in the figure. To illustrate that agenda for action and action itself are in principle different instances – agenda for action is a theoretical construct, while action is a physical and observable phenomenon – a different shape (a diamond) and a different filling colour (grey) have been used. The concept of 'agenda for action' also links the agenda setting model to prospective collective memories. To demonstrate that these scholarly traditions neither stem from each other nor overlap, the rectangles that correspond to collective memories should be seen as belonging to a different dimension, shown by solid border lines. Collective prospective memories, on the other hand, belong to cultural memories. For the sake of simplicity, cultural memories are depicted as an entity that does not belong to theoretical concepts, similarly to action (such that they are filled with the same colour in the figure), though differs from it (hence the different shape for action). Pragmatics and framing correspond to the sources of the second (an expression of a need to act) and the third (a desired course of action) elements of agenda for action denoted by blue and red circles, respectively. The connection between agenda for action and collective action framing as a subfield of framing can be seen from the overlaps between respective zones. One should note that overlaps and connections suggest the linkage of concepts and theories and therefore should not misread to reflect a hierarchy of approaches.

A new concept 'agenda for action' introduced in this chapter can bring the often overlooked motivational component back in the sight of framing scholars. As once pointed out by Steinberg (1999): "…frames offer a diagnosis and prognosis of a problem and a call to action for its resolution" (p. 737). Developing action agenda setting will advance and contribute to the agenda setting theory more broadly. It can also help to draw a clearer border between the issue to be acted upon and the recommended action itself, which agenda setting theory has thus far been unable to provide (Hardy & Sevenans, 2015; Helfer & Wonneberger, 2015). Finally, apart from theoretical gains and advances, the concept 'agenda for action' helps scholars to understand how societies mobilize collective action as well as to predict future developments and happenings. Speech act theory can also benefit from incorporating the idea of 'agenda for action' into its toolkit as it offers a means of uniting propositions that possess a word to world direction of fit. Agenda for action enables the identification of such statements, not only based on speech act verbs, but, also as they extend the theoretical scope through consideration of implicit calls to action.

2 Automated text analysis

To call for action is one of the basic language functions (Jakobson, 1960), therefore agendas for action may be found in almost any spoken or written text. Identification of agendas for action is of vital importance: It enables the speaker to fulfil their communication intent and the hearer to act upon it. Moreover, knowing the object of the call may help to foresee, react to, or moderate behavioural outcomes. Public agendas targeting large groups of people can potentially provoke mass actions and are particularly important to monitor. It is especially important if they concern war and violent conflict, for knowing what actions are being requested by different groups, and which agendas have been accepted and disseminated can help to predict phases of escalation and de-escalation, while tracking the dynamics of the conflict. These agendas are usually communicated via different channels (such as, traditional and social media, political and strategic communication) and in different languages. The same channel can communicate multiple, even contesting, agendas. In order to understand agendas promoted in, say, Russian, English and Chinese media, one has to either be able to understand all those languages, or to use translation services. These avenues make a large volume of heterogeneous data available to scholars of textual analysis, the handling of which in a traditionally manual manner is extremely time consuming. It is practically inefficient to manage this data by hand, for this reason, and to be able to keep track of circulating agendas for action in a timely manner, one has to employ automatic tools for text analysis.

2.1 Introduction to machine learning

The penetration of big data into all areas of science has brought with it many opportunities and challenges. Big data provides insights and makes information about various aspects of life easily accessible. On the other hand, huge volumes of data provide not only useful and meaningful knowledge, but is often full of noise and masses of irrelevant information. Therefore, the biggest challenge of big data is its interpretation and use. The user of big data must filter meaningless noise from the insightful and useful content, and must then interpret the useful information. Moreover, the amount of data available for analysis grows exponentially. It has become clear that traditional manual approaches for big data analysis fall incredibly short, as the time required for the manual analysis of terabytes of information and the hidden dependencies and patterns continue to elude the human mind.

For decades, machine learning has been at the forefront of various fields of science that engaged in analysis of big data. Informatics, engineering, biomedicine, genetics, data science, and computational linguistics, for instance, have used machine learning to successfully analyse and interpret large volumes of data. In communication science, however, machine learning has not yet received due attention (Burscher, Odijk, Vliegenthart, Rijke, & Vreese, 2014; Scharkow, 2012, 2013; Vargo, Guo, McCombs, & Shaw, 2014). Given that big data has increasingly become an important source of information for communication science, it would benefit the field to import the advances and knowledge gained in machine learning from other disciplines, adapting them to the needs of the field.

Being an interdisciplinary work, one of the aims of this thesis is to bring computational linguistics and communication science together by enriching the methodological toolkit of the latter with the methods and techniques of the former. In the course of the present work I demonstrate that machine learning can be used in communications and that it can advance automated content analysis by extracting agendas for action from texts. The work deals with textual data (unlike numerical data, such as, the number of survey respondents, or categorical data, as in the level of education of respondents). In data science, textual data belongs to the class of unstructured data, which is in contrast to the structured data of survey results or library catalogues. Text can be of any genre, style, and topic, containing various syntactic and lexical patterns, with differing grammatical structure. The information that needs to be extracted (agenda for action in our case) can appear in any place in a text and in any form, if at all. In contrast, structured data, such as survey results are far more organised and predictable. All subjects are usually asked the same questions, the variety of answers is far less (usually from a limited pool of options), and results can be analysed and compared against a pre-defined set of criteria. The results of structured data analysis are therefore more easily comparable, as extraneous data are usually not present (i.e. the respondents only provide information that is being asked, they usually do not anything beyond pre-set questions which makes the results easily comparable). All the difficulties of dealing with textual data have enabled computational linguists to come up with specific tricks and techniques for processing and analysis of text. These are discussed in greater detail below.

2.1.1 Definition of machine learning

The seminal definition of machine learning (ML) was given in 1959 by Arthur Samuel. He defined ML as the "field of study that gives computers the ability to learn without being explicitly programmed" (Samuel, 1959, p. 210). A more recent definition of ML is as follows, it is a branch of computer science, the objective of which is to make a computer learn and make decisions based on its own experience, i.e. through data that it has been provided (Kohavi & Provost, 1998). This aim is achieved via developing and training statistical models (referred to simply as 'model' in this work) to recognize patterns in the new data based on the information learned from the previously seen data (training set).

Data may be represented as a set of variables with values and corresponding labels. The computer is expected to assign the same labels to a new item – this process is called supervised learning. It is widely used for text classification, authorship attribution, and sentiment analysis, among others. Identification of agenda for action also falls within the scope of text classification tasks. In the assignment of agenda for action, the computer is given a set of sentences that are already annotated as an agenda for action or not an agenda for action, a ML algorithm is then expected to classify new sentences that have not been annotated yet according to the annotated schema.

If the training data are not labelled and the researcher is exploring the dataset for hidden patterns and dependencies, an unsupervised learning approach could be used. In this approach, the machine will attempt to find subsets of data that are more similar to one another and cluster them together. Using this method, one can automatically cluster texts that belong to the same topic or are written in one language. In our case, it is known what types of information we are looking for – agenda for action – the labels are clearly defined, and the machine is not expected to figure out the data structure. For this reason, I focus on supervised learning as the method used in present thesis.

2.1.2 Training corpus

In order to train a supervised learning model, the first step is to label data. Each data item is tagged with a category to which it belongs. If the task is to train the model to automatically identify the topic of a given text, then each data item (a text) in a training corpus should be annotated with a topic, or a genre, if the learning objective is to identify genres. If the algorithm is expected to automatically measure the sentiment of a document (either the whole text, or one paragraph, or just one sentence), then each data item in the training corpus needs to be assigned a sentiment value.

Data labelling is done by human annotators. First, clear labelling guidelines must be developed and the annotators need to undergo training. It is good practice for authors of a study not to take part in labelling, so as to avoid bias. Inter-annotator agreement is usually measured during process of labelling, with higher agreement signifying better quality, as the guidelines are clear and the information that is being labelled is unambiguous.

There exist multiple labelled corpora for different tasks. Reuters collection of news (Lewis, Yang, Rose, & Li, 2004) is often used for text categorisation. There is also a corpus of news frames (Card, Boydstun, Gross, Resnik, & Smith, 2015), and others ('Kaggle', 2015). While the former is just a set of texts grouped in accordance with topic, and the latter classifies larger parts of texts on a rather abstracted high-level (for example, a classification may be, the first paragraph presents the frame 'economy', the last paragraph presents the frame 'unemployment'), neither corpus includes annotation of agendas for action. There are several speech act corpora (Stolcke et al., 2000; Leech & Weisser, 2003) which could better serve the purpose of this thesis, but these are used for recorded speech and are not suitable for written texts. Moreover, they classify speech acts as requests, orders, inquiries, plans, among others, and do not for the particular requested action. Therefore, none of the abovementioned corpora can be used for the task at hand and the new corpus for agenda for action classification must be developed.

2.1.3 Rendering data into features

When one says that algorithm learns from data, one means that for a machine to be able to draw conclusions from data, data need to be pre-processed, rendered into a machine-readable format. Dealing with textual data is a fairly hard task. Words and word combinations that make sense for humans are a meaningless sequence of code for a machine. Strings of text need to be represented as figures and scores for computer to make sense out of them. In other words, 'features' need to be extracted from data for the machine to use it as learning material. Ideally, these features reflect properties or features of the text that are indicative of the classification task at hand; such that different classes of texts or expressions differ from one another regarding these features. For authorship attribution, for instance, sentence length, the number of clauses in a sentence, punctuation patterns, and specific word usage, have been used as features. The set of features used to train the algorithm determines the performance of the model to a great extent.

In natural language processing the most common approach to text classification is using *n-grams* as features (Cavnar & John M. Trenkle, 1994; Fürnkranz, 1998). With n=1, or unigrams, one deals with a 'bag-of-words', and each word (whether it occurs and how many times) is treated as a feature in itself. With n>1, two, three or n neighbouring words are considered as one unit, or one feature. The example sentence 'Peace is the only answer' will be presented as the following set of bigrams: 'Peace is', 'is the', 'the only' and 'only answer'. With trigram representation it will look as follows: 'Peace is the', 'is the only', 'the only answer'. The appropriate size of n is usually determined experimentally. The bag-of-words usually misses a lot of important connections between words both semantical and grammatical. For example, 'White House', if split into unigrams, loses its meaning as the political body. Similarly, the word 'need', in isolation, does not tell whether it is a noun as in 'an urgent need' or a verb as in 'need to proceed'. If the size of *n* increases significantly, then the number of occurrences of each unit is equal to one in most cases. When this occurs, lexical and grammatical patterns are lost, which renders statistical analysis meaningless. It is good practice to have a range of possible *n*, usually between 1 and 5. Such an approach enables the scholar not only to account for single words, but also to consider word collocations, and some longer pieces of text that might be re-occurring.

In most scenarios the text samples used in the training corpus are of different length; meaning that a word may have a higher occurrence score simply due to the greater text length. This property, however, does not mean that it contributes more to understanding the meaning of the text. If a word occurs five times in a text of one hundred words, it may still be more important for the core meaning, than if it had occurred ten times in a text of five hundred words. It means that a word with the occurrence score five should be given more weight during classification routine than a word scored ten.

In order to handle this problem, normalization is usually performed. One of the standard normalization procedures is to use term frequencies (TF) instead of term occurrences. TF can be computed by dividing the number of occurrences of a given word by the total number of words in the text, as shown in Equation 1.

$$TF(t) = \frac{N(t)}{N}$$

Equation 1: Term frequency computation

Where *N(t)* is the number of occurrences of the term *t*, *N* – total number of words in a sentence

Due to language peculiarities, some words tend to occur more often than others. Such is true for grammatical and functional words such as prepositions and articles. Likewise, more general words occur more frequently than more specific ones (for instance, 'weapon' is a more frequently used word than 'bazooka'). For text classification though, the more seldom words are usually more important. Knowing that 'the' and 'or' are the most frequent words in a text does not help to identify the subject of the text. Likewise, even the presence of the word 'weapon' does not necessarily render the text with the meaning of war, while 'bazooka' would be a strong indicator that the text belongs to the 'war' topic. To account for such words that frequently occur in many documents, those are not very informative for the purposes of classification, the following procedure is normally used. TF is multiplied by inverse document frequency (IDF), a logarithm of the total number of documents in the corpus divided by the number of documents where the term occurs (Jones, 2004; Salton & Buckley, 1988; Salton & McGill, 1986; Wu, Luk, Wong, & Kwok, 2008). The resulting number is known as a TF-IDF score and it can be computed for any *n*-gram. TF-IDF is widely used as a feature in many natural language processing (NLP) tasks including text classification. It can be derived using Equation 2:

$$\frac{N(t)}{N} * \log\left(\frac{N(s)}{N(s_t)}\right)$$

Equation 2: TF-IDF computation

Where $N(t)$ is the number of occurrences of the term t, N – total number of words in a sentence, $N(s)$ – total number of sentences, $N(s_t)$ – total number of sentences where a term t occurs.

Another approach for dealing with high-frequency words is to use a stop-words list (Silva & Ribeiro, 2003). Such a list normally includes auxiliary words with limited semantic information, as in those that mainly fulfil grammatical functions. A stop-words lists may include, for example, 'a', 'the', 'is', 'was', 'and', 'or', 'of', 'as'. This approach is a complementary to the TF-IDF described above. Stop-words removal excludes words that do not bare semantical meaning from the analysis. TF-IDF up-weights the words that intuitively contribute more to sense disambiguation. Although, stop-words removal is a very common step, it sometimes can worsen the analytical result. To classify the texts with respect to their topic, articles and prepositions can be removed harmlessly; however if the text is rather short and grammatical information is important for classification, as in speech act classification, functional words are usually the carriers of grammar information. Their removal can harm the performance of automated analysis. The usual trade-off here is to tailor the stop-words list with regard to the task at hand (Silva & Ribeiro, 2003; Zaman, Matsakis, & Brown, 2011).

Another pre-processing step in the analysis of textual data is lemmatization or stemming. The former is bringing all morphological variations of a word to its canonical form. Plural numbers are reduced to the singular. All tenses and verb forms are reduced to the infinitive. Comparative and superlative degrees of adjectives are reduced to the positive. For instance, 'am', 'being', 'were' are rationalised to 'be', or 'document' and 'documents' to 'document', or 'bigger', 'the biggest' to 'big'. Similarly, stemming consolidates a word to its stem form. For example, 'terrorism', 'terrorist', 'terrorize' would become 'terror' (Balakrishnan & Ethel, 2014). Analogous to stop-words removal, performing either of these steps can be beneficial to large text classification, especially when the grammatical information carried by removed suffices does not impact textual meaning. When the classification task involves grammatical information, stemming or lemmatization can reduce the efficacy of classification in a similar manner to the stop-words removal and TF-IDF processes.

2.1.4 ML Algorithms

Once the data are prepared and the features are extracted, they can be fed to a ML algorithm, which completes the analysis. The process of feeding the data to the algorithm and having it learn the patterns is usually referred to as training. We say that the model is being trained to later make predictions based on new data.

There exists a great variety of learning algorithms. The selection of algorithm usually depends on the task at hand, amount of data available, computational resources, among other key factors (Abu-Mostafa, Magdon-Ismail, & Lin, 2012; Conway & White, 2012; Hastie, Friedman & Tibshirani, 2001; Segaran, 2007).

The algorithms can be grouped based on the principle they work according to:

- Regression algorithms model the relationship between variables that is iteratively refined using a measure of error in the predictions made by the model. The most common representatives of this algorithm family are linear and logistic regression, widely used for classification and value prediction.

- Tree based algorithms construct a model of decisions made based on actual values of attributes in the data. The purpose of the algorithm is to find the shortest path from the top node to the bottom. The algorithms of this family mimic the way humans make decisions. The results are relatively easy to interpret, and the models are very powerful, capable of handling both linear and non-linear problems. The decision tree is an example algorithm.
- Bayesian algorithms are probabilistic models that apply Bayes' Theorem[3] for problems such as classification and regression. They output the probability that a data item belongs to a given class. Naïve Bayes is the most popular representative of this class.
- Clustering algorithms group the subsets of the whole data set together based on the distance between data points. *k*-means and *k* nearest neighbours are the most typical example algorithms.
- Margin classifiers are classification algorithms that draw a separation line in cases of linear classification, or a hyperplane in cases of multidimensional classification. Separation is based on the distance measurement, in such a way that each data point is maximally remote from the separation margin. The support vector machine is the best-known algorithm from this group.
- Sequence modelling algorithms, unlike other classifiers that predict a single output label ignoring the contexts, are able to make structured predictions based on context, assigning output labels to sequences of data points. Conditional random fields is an example algorithm from this class. It is widely used for part-of-speech tagging, grammar parsing and machine translation.
- Artificial neural networks – or the more advanced version, known as deep learning algorithms (Goodfellow, Bengio, & Courville, 2016) – are learning algorithms inspired by processes in the human brain. Similar to biological brain, wherein information is processed by complex networks of axons and dendrites passing information as electrical current, artificial neural networks consist of multiple layers of artificial neurons. The first layer, known as an input layer,

3 Bayes' theorem describes the probability of an event given that the conditions related to the event are known (Bayes & Price, 1763).

takes in the data, the last layer outputs the result labels or values. All the layers between the input and output layer are called hidden layers. This group of methods derived its name, deep learning, from the hidden layers. The number of layers is arbitrary and is defined experimentally. The main objective of a deep artificial neural network is to learn multiple levels of composition. Each layer receives the calculation results from the previous layer as its input, performs some mathematical transformation, and sends the result to the next level. In such a manner, the whole training set passes through all layers of neurons. Each layer generates a new representation of data which is hidden from the user, which is exposed only to the result of the output layer, the desired classification label, for example.

Agenda for action extraction and categorisation is a classification task. In order to decide which classification algorithm to use, several experiments need to be run. The algorithms used in the present work and their in-depth analysis will follow, as I describe the task of agenda for action extraction and classification.

2.1.5 Evaluation of the performance of ML algorithms

Data are crucial to any ML pipeline and are needed at several points in the process. One subset of data is needed to evaluate different algorithms during the development phase, assisting in the selection of the most promising one to be used in the application. Another chunk of data will be used for training the model, and a third chunk serves to test the performance of the model, a chunk of the dataset has to be saved for the testing phase. In practice, development is often carried out on parts of training subsets, but a reserved set is required for validation. If the model was shown a data item during training and then is given the same item during testing, it will reproduce the output label that it has memorised, rather than learning which patterns and dependencies in the data structure determine its assignment.

In other words, all data primarily serve three purposes: model development, model training and model testing. Hence, the dataset should be split into three respective subsets. The traditional split, albeit arbitrary, would be 20–30% for development, 50–60% for training and 20–30% for testing (Dobbin & Simon, 2011; Reitermanova, 2010). If data are limited, *n*-fold cross-validation may be used. In this process, all data are split into *n* folds. For each iteration, a fold is left out to be used for testing and the remaining *n*-1 folds are used for training. *N* iterations are performed in total (Stone, 1974). The performance of the model is evaluated based on the accuracy, precision, recall and f1 scores calculated on the test set. In case of *n*-fold cross validation, the mean over all *n* iterations is computed.

Accuracy is the number of all correctly classified instances divided by the number of all instances (Hall et al., 2009). The equation for its calculation is present below:

$$accuracy = \frac{\mathrm{tp} + \mathrm{tn}}{\mathrm{tp} + \mathrm{tn} + \mathrm{fp} + \mathrm{fn}}$$

Equation 3: Accuracy computation

Where *tp* stands for true positive, *tn* – for true negative, *fp* – for false positive and *fn* – for false negative

However, when dealing with highly imbalanced classes, this metric does not suffice, as all instances would be coded with the label of the most frequent class, and in so doing, all the instances of the underrepresented class would be classified falsely. Even in these cases, the accuracy score would still be high. To account for this problem, precision and recall are commonly employed to measure the performance of ML algorithms. Precision determines how many of the instances classified as class A indeed belong to the class. It is a fraction of true positives and a sum of true and false positives and can be computed with Equation 4:

$$precision = \frac{\mathrm{tp}}{\mathrm{tp} + \mathrm{fp}}$$

Equation 4: Precision computation

Recall measures how many instances of class A are classified as such and is computed by division of true positives by the sum of true positives and false negatives (Witten, Moffat, & Bell, 1999), as shown in Equation 5.

$$recall = \frac{\mathrm{tp}}{\mathrm{tp} + \mathrm{fn}}$$

Equation 5: Recall computation

Neither precision nor recall, as stand-alone metrics, give enough information about the quality of the classifier. A model can correctly identify the instances of A (for example, when no instance of class B is labelled as A) and thus have high precision, but it can skip most of the A instances (having low recall). Similarly, a model can find all the instances of A (having high recall), but also label as such many instances of B (low precision). To solve this problem, the f1 score has been developed. It is the harmonic mean of precision and recall, computed by dividing the product of precision and recall by their sum (van Rijsbergen, 1979), as shown in Equation 6.

$$f1 = \frac{\mathrm{precision} \cdot \mathrm{recall}}{\mathrm{precision} + \mathrm{recall}}$$

Equation 6: f1 computation

2.2 Overview of existing techniques to extract directive statements

As shown in the first chapter, the idea of 'agenda for action' has been approached by multiple disciplines but have not yet been the sole focus of any scholarly tradition. The same holds for automatic extraction of agendas for action. Multiple fields have developed approaches that touch upon agenda-for-action-like information identification, but in fact do not extract agendas for action from texts, nor classify them. Below is an overview of relevant work in communication science and computational linguistics that approaches the problem from different perspectives.

2.2.1 Communication science approach

In communication science, works that aim at identifying treatment recommendations are the most relevant to agenda for action extraction. However, these studies deal with treatment recommendations as one of the frame elements. Agenda-for-action-like entities are extracted from texts with the purpose of capturing the frame. There is no interest in the treatment recommendations as such. Matthes and Kohring (2008), who suggested coding each frame element as defined by Entman (1993) in order to measure it, have been the pioneers of this approach. Thereafter several case studies have followed that targeted very specific topic-dependent treatment recommendations. For instance, in the study of the news coverage of nanotechnology, Donk, Metag, Kohring, and Marcinkowski (2012) operationalised treatment recommendation as the call for regulation or support of nanotechnologies. Bowe, Oshita, Terracina-Hartman, and Chao (2014) in their analysis of climate change coverage, have coded treatment recommendations from the perspective of whether the issue had been depicted as a real phenomenon requiring action, or the false one. In the study of the coverage of mosque construction in the US, Bowe (2014) has coded treatment recommendation as a call in favour of, or against building a mosque.

Many scholars have utilized computer assisted methods to measure frames. Automated frame analysis has focused mostly on the operationalisation of a frame as the main organizing idea, which can be captured based on specific lexical co-occurrence patterns (Al-Rawi, 2015; Arrese & Vara-Miguel, 2015; Baden, 2010; Kutter & Kantner, 2011; van Atteveldt, 2008). In these works, treatment recommendations have not been a feature.

Sanfilippo et al. (2008) have approached automatic frame extraction from a different angle, namely they captured frame elements such as Promoter, Target, Issue, etc., treating frames as a rather linguistic phenomenon (Fillmore, 1976). One of the measured frame elements was *Intent* measured by capturing verbs with respective meaning (e.g., 'intend', 'plan', 'be going'). From the viewpoint of pragmatics, these verbs perform an action by virtue of their verbalisation, thus they render the utterance into a speech act. In accordance with Searle's (1975a) speech act classification, these are grouped as commissive speech acts. Given that commissive speech acts are one of the ways to express agendas for action, this study appears to be the closest to the present thesis, in the sense of automatic capturing of agendas for action in text. Despite this, intent was identified only as a frame element in this study and not as an independent entity.

Studies from the 2000s have used methods of statistical learning in the domain of communication science. A foundational study was published by Vargo et al. (2014), who used machine learning for sentiment analysis of tweets during the 2012 United States presidential campaign. Tsur, Calacci, and Lazer (2015) have used unsupervised methods of topic modelling and time series to identify topics, political agendas (temporal shifting and ownership) used by politicians to run political campaigns, based on the statements of the members of Congress in the US. Greene and Cross (2016) have used topic modelling to track agenda changes in the debates of the EU parliament.

2.2.2 Computational linguistics approach

2.2.2.1 Short text classification

Computational linguistics is a branch of science which engages extensively with automatic text analysis, usually information extraction and text classification. As agendas for action take the form of a textual proposition, the task of agenda for action extraction and classification falls in the scope of short text classification. As has been pointed out, due to the length and origin, short text usually appears to be noisy and sparse data, for which the traditional bag-of-words approach does not suffice. When the bag-of-words representation is used, most of the words for each sentence will have TF-IDF, occurrence or frequency scores of zero. Those that have positive scores, will rarely exceed one (Rosso, Errecalde, & Pinto, 2013; Song, Ye, Du, Huang, & Bie, 2014). Methods to overcome these difficulties have become the core interest in the area of computational linguistics. Much of the scholarship to 2019 has addressed this problem.

A number of scientists have experimented with different learning algorithms, including deep learning, and have fine-tuned parameters to meet the challenge (Ali, Khalid, Rana, & Azhar, 2018; Khoo, Marom, & Albrecht, 2006; Kim, 2014; Revathi, Ramya, Tanuja, Pavani, & Swathi, 2012; Xu, Sun, Deng, & Tan, 2017). A lot of work has been done in the field of feature engineering, for example in feature space extension. Instead of extracting features from the given corpus only, which most likely possesses a very limited number of words due to length, the features are enriched with the scores computed for related longer texts drawn from external knowledge bases such as WordNet or Wikipedia (Hu, Sun, Zhang, & Chua, 2009; Phan, Nguyen, & Horiguchi, 2008; Schonhofen, 2006). This approach is often combined with latent semantic analysis, enabling the representation of a document in such a way that others with semantically related terms are located close to each other, thereby increasing the likelihood of assigning the same label (Hofmann, 1999; Landaues, Foltz, & Laham, 1998; Phan et al., 2011; Pu & Yang, 2006; Sahlgren & Cöster, 2004; Zelikovitz & Marquez, 2005).

Another way to extend the feature space is to use meta-data (such as user age or hashtags) as features. This approach shows good results when classifying tweets as an event, news piece, or private message (Sriram, Fuhry, Demir, Ferhatosmanoglu, & Demirbas, 2010). Scientists have successfully used distributional term representation (DTR) for short text categorization. DTR allows presentation of a document by document-occurrence and term-co-occurrence statistics, resulting in a short-text represented by the combination of the contexts of its terms. In so doing, the sparseness is reduced and the issue of terms with low frequencies is addressed (Cabrera, Escalante, & Montes-y-Gomez, 2013).

An alternative solution is document expansion, achieved by accounting for weights associated with terms that do not occur in the document, but are associated with terms that do occur (Fan & Hu, 2010; Nagarajan et al., 2007; Wang, Zhou, Li, Hu, & Hu, 2009; Yan & Wang, 2010). Performing a high-frequency feature extension, based on the latent Dirichlet allocation (LDA) topic model (Hu, Jiang, & Chang, 2013; Pinto, Jiménez-Salazar, & Rosso, 2006), may also address the issue. As mentioned above, a short text is sparse data, important words that contribute to classification might not appear, thus their score is zero (low frequency feature). Other words that usually occur often (as articles and prepositions) do not help with classification. LDA is a topic model which enables the identification of high frequency features – words which, according to the topic of the document, should have higher frequency. Thereafter these features are extended, zero or extremely low scores are adjusted in accordance with the document topic.

Finally, for thematic classification, external data may be used. The model is trained per normal methods, but when it is given a new instance to classify, it can refer to an external text which may belong to the same topic of the training data instance. For example, in order to classify a new data point as a text about politics, one may look to all texts about politics available on the Web. Similarly, the instance in question can be connected to an external resource. In such cases as when the training set consists of article headlines to classify other article titles, one can use the body of the article for reference (Nigam, Mccallum, Thrun, & Mitchell, 2000; Ramírez-de-la-Rosa, Montes-y-Gómez, Solorio, & Villaseñor-Pineda, 2013; Zelikovitz & Hirsh, 2000). In order to carry out these approaches, similarity measurements are needed to measure the degree of similarity between the current data instance and the external text. This measurement involves additional computations.

A number of experiments in sentence classification have been conducted in the medical domain assigning such labels as outcome, diagnosis, and treatment, as well as investigating additional features to improve the performance (Kim, Martinez, & Cavedon, 2011; McKnight & Srinivasan, 2003).

2.2.2.2 Automatic classification of speech acts

Another component of the computational linguistics literature that is highly relevant for this thesis investigates machine learning approaches that classify text segments as speech acts (manually annotated speech act corpora are used). Cohen and Carvalho with colleagues categorize whole email messages as requests, proposals, amendments, commitments, deliveries, and other speech acts (Carvalho & Cohen, 2005, 2006; Cohen, Carvalho, & Mitchell, 2004). Using TF-IDF-weighted bag of words, bigrams, and POS-tags, they compare four classifiers – Voted Perceptron, AdaBoost, Support vector machine (SVM), and Decision Tree – the last two outperforming the rest. Tavafi, Mehdad, Joty, Carenini, and Ng (2013) continued the work by developing a domain-independent classification strategy. They used unigram frequencies together with the length of the utterance to train SVM and Conditional Random Field classifiers, with the latter outperforming the former in all the experimental setups.

Qadir and Riloff (2011) moved the categorization task to sentence level, assigning speech act labels, as defined by Searle (1976), to message board posts. In their classification model the researchers also accounted for the grammatical structure of sentences by capturing imperative sentences and discerning them from interrogative one, while using the bag-of-words. Oya and Carenini (2014) integrated the task of speech act identification with email summarisation. To identify speech acts they used *n*-grams together with POS information, performing additional routines to reduce sparseness. They used Dynamic Conditional Random Fields to perform complex structured classification and prediction.

Jeong, Lin, and Lee (2009) applied a semi-supervised method for speech act classification in email and forum threads. They used several annotated datasets, performing domain adaptation routines to fit unlabelled data, afterwards applying bootstrapping techniques to identify speech acts. They used dependency trees[4] as features.

Shafiq, Carenini, and Lin (2011) proposed unsupervised dialogue act modelling for email and forum threads. They introduced graph-based and two probabilistic unsupervised approaches for modelling dialogue acts. By comparing those approaches, the researchers demonstrated that the probabilistic approaches were quite effective, performing better than the graph-based method.

Joty and Hoque (2016) applied deep learning for speech act identification using a recurrent neural network[5] to classify sentences in asynchronous conversations as speech acts. They categorise the speech acts as statements, requests, questions, suggestions, and others.

Wood, Rodeghero, Armaly, and McMillan (2018) have extracted and classified speech acts from the dialogues of software developers with a virtual bug fixing assistant. To that end, they have recorded the dialogues, annotated the corpus and applied Naïve Bayes, SVM and logistic regression algorithms. Logistic regression has shown the best classification results.

4 Dependency tree is a linguistic representation of the sentence structure: Traditionally, a predicate is the "head" of a sentence. It subordinates nouns and adverbs; nouns, in turn, are linked to adjectives, prepositions and other nouns. These links are called dependencies. See Tesnière (1976 [1959]) for more details.

5 A recurrent neural network is an example of a deep learning algorithm which is able to make conclusions not only based on the input at the given point of time, but it also takes the history (the parameters it has already learned) into account (Hopfield, 1982).

2.3 Conclusion

Machine learning as a method of text analysis has been the main approach in computational linguistics and has gained popularity in other fields of science, including communication science. One of the biggest advantages of machine learning for content analysis is the ability to make use of big data, rendering analysis data-driven, which helps to eliminate human bias and saves resources.

There exist approaches and solutions to extract and classify speech acts, directive statements and treatment recommendations as one of the possible subtasks of text analysis. However, each of them has been operationalized within a single scientific tradition. Speech acts and directive statements have been operationalised within computational linguistics, while treatment recommendations have been operationalised within communication science. Thus, their analytical strategies are divergent. The mainstreaming of automatic processing efforts can benefit each of these scholarly traditions. In the next chapter, I show that these approaches can be merged and, in line with advances in each approach, suggest an automated method to extract and classify agendas for action.

3 Automatic extraction and classification of agendas for action

In the following chapter, I explain the manner in which the ideas and principles outlined above have been used and amended in the present thesis to extract and classify agendas for action.

3.1 Agenda for action corpus

The content to be extracted, agendas for action, is well-defined. The categories into which the data will be classified, per the agenda for action taxonomy, is also known. Agenda for action extraction fits the framework of supervised learning defined above. Therefore, before commencing a classification routine, a training corpus needs to be crafted and annotated. Despite ample interest in agenda for action across different disciplines, to the best of my knowledge, there is no pre-defined corpus suitable for this work ('Kaggle', 2015). For this reason, I first had to prepare a training corpus.

As mentioned in Section 1.1, the present dissertation has been conducted in affiliation with the FP7 EU-Project INFOCORE[6]. One of the main purposes of the project was to investigate how violent conflict is presented by media. The research consortium was studying traditional printed media, social media, namely Facebook and Twitter, strategic communication materials circulated by public figures and NGO's and materials produced by political actors. The affiliation with INFOCORE determined the choice of data sources for agenda for action corpus. Due to INFOCORE's thematic focus, the search terms for texts to be included in the corpus were 'war', 'violence', 'crisis', 'conflict' or 'peace'.

To cover the whole range of mediated communication, four data sources have been used. Traditional media is represented by newspapers. In order to construct the newspaper corpus, all English-language sources stored on LexisNexis[7], and published between 01.01.2015 and 01.03.2015, were selected.

6 www.infocore.eu, Grant Agreement No. 613308
7 https://www.lexisnexis.de

Strategic communication sample consists of materials published by non-governmental organizations (NGOs), military organizations, and social movements. These are texts in the form of news articles, reports, press releases, and interviews. For this study, I chose texts produced by the following actors: NATO, UNICEF, UNESCO, UNDP, UNHCR, Red Cross, and Human Rights Watch. All articles from these sources were published between 01.05.2015 and 31.12.2015. The reason to choose these organizations was the field they operate in and accessibility of data.

The political communication subset is represented by the transcripts of the parliamentary debates in the House of Commons in the UK in 2013 and by adapted texts from the EU Parliament published in the range 01.06.2015 till 01.05.2016. These sources were chosen due to the ease of data access and due to the fact that they are produced in English.

Finally, I also included social media contents, such as found on Twitter in particular. Other social media platforms were excluded from this study because of the difficulties associated with obtaining data from them or the lack of textual data as in case of Instagram. Tweets containing the same search terms, or hashtags, published on 01.02.2016 were chosen for the corpus.

As mentioned, computational analysis of text requires the subject to be represented as a set of discreet data points. Assuming that an agenda for action takes the form of a proposition, which in natural language is usually expressed with a sentence, then a sentence is the unit for analysis in the present work. Hence, after being retrieved, all the texts were split into sentences. The split criterion was terminal punctuation (a full-stop, question and exclamation marks, ellipsis, the beginning of a new line and the beginning of a new paragraph). Thereafter, a random subset of sentences was taken from each discourse, for uniform representation. The sentences were shuffled and each sentence tagged as one expressing an agenda for action or not. Those expressing an agenda for action were classified in accordance with the developed taxonomy (see Section 3.8.2).

Three coders took part in corpus annotation. One hundred sentences were used to measure the inter-coder agreement. Krippendorff's α (Krippendorff, 2004a, 2004b) was used as the metric for inter-coder agreement. Krippendorff's α is recognized as one of the most suitable and widely used methods for content analysis in communication science and in computational linguistics, when more than two coders take part in annotation, interval metric is used, and the amount of data for validation is small (Artstein & Poesio, 2008; Gwet, 2014; Hayes & Krippendorff, 2007; Lombard, Snyder-Duch, & Bracken, 2002; Neuendorf, 2002). The Krippendorff's α for the two labels (agenda for action versus not agenda for action) was 0.642.

Communication science and computational linguistics interpret the agreement score differently. While communication science takes a rather strict approach and considers the alpha above 0.8 to be significant (Krippendorff, 2004a), computational linguistics lacks a unified approach, with scores above 0.4 regarded as adequate (Landis & Koch, 1977; Marion, 2004). Moreover, when annotated corpora are used for training machine learning models, the agreement score does not necessarily affect algorithm performance. Even if the measured agreement between coders is high, but those instances that the coders disagreed upon contain patterns (in other words, coders disagree consistently), the results of machine learning will be rather poor because those patterns where coders disagree will also be learned by the algorithm and those instances will be consistently misclassified (Reidsma & Carletta, 2008). The inter-coder agreement score for the corpus developed for this thesis may be considered insufficient from communication science perspective. However, given that the annotation task was of a linguistic nature (targeting grammatical and semantic patterns), and that the corpus will be used to train a statistical classifier, I follow computational linguistics, treating the score as sufficient for the purpose of the current research. It is worth noting that treatment recommendation is one of the most challenging frame elements to be coded (comparing to issue definition, moral evaluation and causal attribution). Bowe et al. (2014) have reported Scott's π of 0.75[8] and Bowe (2014) has reported a Krippendorff's α of 0.64 for the given element.

8 Scott's π is an index to measure inter-coder agreement. It takes the number of categories and values distribution among them into account. Its main shortcoming is the assumption that the coders distribute values across categories identically. The index is not suitable for nominal values and when more than two coders participate in coding (Scott, 1955). According to Neuendorf (2002), Scott's π below 0.8 signals of great disagreement. Other scholars suggest that for Scott's π more liberal criteria are acceptable (Lombard et al., 2002).

The resulting corpus consists of 14,876 sentences in total, each annotated as expressing an agenda for action or not. The corpus is balanced. Agendas stemming from each source comprise roughly a quarter of all sentences. The share of agendas for action among texts is not equal. More than a half of the sentences in the political sample express an agenda for action (which makes sense as the main objective of political communication is to advance policies, and thus prescribe actions). The share of agenda for action within journalistic and strategic texts is between 21% and 27%, with only 15% of tweets calling for action. The corpus structure is presented in Table 3:

	Agenda for action	**Non-agenda for action**	**Agenda for action (%)**	**Total sentences:**
Traditional media	920	2464	27.18	3384
Strategic communication	853	3145	21.34	3998
Political communication	2568	1178	68.55	3746
Social media	576	3172	15.37	3748
Total:	4917	9959	33.05	14876

Table 3: Corpus structure

3.2 Features indicative for agendas for action extraction

As described in Section 2.1.3, when dealing with textual data they need to be represented as machine readable features, which are used to train a ML algorithm. For agenda for action extraction and classification, three types of features were used; TF-IDF weighted *n*-grams, linguistic features, and data source.

3.2.1 TF-IDF weighted *n*-grams

For the present work TF-IDF weighted *n*-grams were used. The value for *n* has been decided empirically. Multiple experiments were run with *n* between 1 and 10, the result showed the best performance scores were achieved with *n* between 1 and 4, hence unigrams, bigrams, trigrams and four-grams have been extracted from the corpus. Thereafter, TF-IDF scores were computed for each, per the method explained in Section 2.1.3. These scores were used by the classifier to predict the class of each sentence.

3.2.2 Linguistic features

From the view of pragmatics, any proposition can express a call to action in a specific extra-linguistic context, especially in spoken text. However, when a text is isolated from the extra-linguistic situation, as in the case of journalistic texts or tweets, certain conventions should be observed and additional criteria met for a sentence to possess an illocutionary force. In most cases it is expressed via force-markers (specific verbs, sentential mood or word combinations), which, in principal, are grammatical or lexical patterns. Hence, there exists a finite number of ways to express agendas for action.

The classification routine presented here consists of two steps of differing nature. On the one hand, deciding whether a sentence expresses an agenda for action not only relies on specific words or word combinations, but also depends heavily on grammatical features of the sentence, such as part of speech, tenses and moods, and grammatical dependencies. On the other hand, classifying agendas for action based on the advanced course of action is, to a large extent, a semantic task, involving the semantic and pragmatic information of the utterance, rather than grammar.

In order to up-weight the role of grammar in the first step of the classification routine, classifying sentences as agenda for action, I developed a relevant set of linguistic features. To do so, a careful analysis of the training corpus was conducted. Certain lexical and grammatical patterns allow determination of whether a sentence is calling for action, and how one might express them formally, were identified. Some of the features rely on word lists (as modal verbs or speech act verbs), some of them have been inspired by similar research (e.g., Khazaei, Lu & Mercer, 2017; Qadir & Riloff, 2011), some are based on intuition of the speaker of a language. Linguistic features are used in addition to TF-IDF weighted bag-of-words and *n*-grams for first-round classification.

Unlike TF-IDF scores, which are numerical features, linguistic features are binary. They can take two values, either 0 (if a sentence does not possess a feature) or 1 (if the sentence possesses a feature). For this reason, the definition of the features reads as a statement that can be answered yes or no.

I have included 32 linguistic features in the analysis.[9] They can be split into several groups, where the features within one group are interdependent and complementary.

Modal verbs

The first group of features addresses the cases when agenda for action is expressed explicitly with modal verbs as in *'They must obey.'* The features of this group do not purely rely on the list of words, but also take into consideration parts of speech and syntactic information. This arrangement filters out those cases in which a word homonymous to a modal verb presents, but, due to its context, does not call for action. In order to disambiguate the words that are homonymous, the term 'candidate' is used. Before it has been established that the word is actually a verb, it is referred to as a 'verb candidate'.

Feature 1: if a modal verb candidate is in the sentence. Here, only four verbs are considered: 'must', 'should', 'need', 'ought', as those may imply a command or a directive statement.[10]

Feature 2: if a modal verb candidate has the part of speech tag of verb (modal or main verb).[11] This feature is supposed to disambiguate cases like *'Their needs have increased'* and *'Something needs to be done'* – the word 'needs' appears in both sentences, the former does not express an agenda for action, while the latter does. Knowing that in the former sentence 'needs' is a noun, and a verb in the latter, may help to disambiguate these cases.

Feature 3: if a modal verb is in auxiliary relation with the verbs 'feel' or 'to be.'[12] An auxiliary of a clause is a subordinate verb of the clause,

9 Stanford CoreNLP toolkit has been used for feature engineering (Manning et al., 2014).
10 The lists of all hint words are presented in Appendix 2.
11 For part of speech tag set details and explanations see Santorini (1990).
12 For grammar dependencies see Silveira et al. (2014).

e.g., a modal auxiliary. This feature is supposed to capture cases as *'It must feel so cold'* where 'must' rather expresses a hunch or assumption rather than committing someone to an action.

Feature 4: if a modal verb is followed by 'have' + past participle. This feature has been introduced to account for sentences, expressing a counterfactual desired or expected outcome, as in the example '*Osama bin Laden should have been captured*.' Sentences like these have been excluded from the scope of agenda for action as they are not addressed to the future (the third property of agenda for action).

Features 2–4 have been developed to complement Feature 1. They depend on it and can only hold, if Feature 1 is true.

Speech act verbs

The next group of features deals with speech act verbs. Commissive speech act verbs such as 'vow', 'guarantee', 'threaten', and directive ones such as 'command', 'call for', 'beg', are considered to signal an agenda for action. Currently, the list includes 62 words (see Appendix 2) and was designed based on the list of speech act verbs suggested by Wierzbicka (1987). The original list was first filtered by the speech act verbs not expressing agendas for action (including partially expressives, declaratives, and representatives). Thereafter, it was further refined by scrutinising the data and analysing classification errors. Some words that were too ambiguous, mostly yielding sentences that do not call to action, were also excluded. For instance, the verb 'push' can be used as a synonym for the verb 'urge', 'command' as in *'They have been pushing the peaceful solution*'. Due to ambiguity, however, most of the sentences that included this verb expressed a different meaning, namely 'to move something away'. Developing linguistic features that would disambiguate these meanings did not improve results. As a result, this verb was excluded from the list of speech act verbs. In future, the list can be further modified and enriched, especially if new disambiguating features are developed. In a similar manner to the first group of features, features 6 to 10 are dependent on feature 5 and can only hold if it is true. Features 7 and 9 address both modal verbs and speech act verbs.

Feature 5: if a speech act verb candidate is in the sentence.

Feature 6: if a speech act verb candidate is tagged as a verb. Like Feature 2, this one is supposed to disambiguate a word that is morphologically identical to a speech act verb but is actually a noun, as in 'User request is being processed.'

Feature 7: if a speech act verb or a modal verb is in an auxiliary relation or has a clausal or prepositional complement (with specific prepositions only). This feature is designed to distinguish between '*He called for peace*' and '*He threatened to start bombing*' and '*He called on the phone.*' the first two sentences express an agenda for action, the third one does not. All three sentences contain speech act verbs ('call' and 'threaten'). To disambiguate the classification of similar sentences, the grammar dependencies must be examined. The second sentence has a clausal complement ('to start bombing'), the first and the third sentences have prepositional complements ('for peace' and 'on the phone' respectively). That is why it is important to consider preposition connecting the verb with its complement: If the verb 'call' is linked to its complement via preposition 'for', then the sentence contains an agenda for action, if it is liked by another preposition – it is not.

Feature 8: if a speech act verb has a direct object. This feature is supposed to catch sentences like '*They want us to go*.'

Feature 9: if a speech act verb, or a modal verb, is tagged as a participle. This feature captures cases, when hint words are modifiers rather than predicates, as in the sentence '*to distribute materials needed by our movement.*' In this example, even though a modal verb appears, it fulfils a modifying function as a clausal complement and does not call for any action.

Feature 10: if a speech act verb is a part of a relative or conditional clause introduced with such relative pronouns as 'who', 'whom' or 'which', conjunctions 'whether' or 'if' and conditional expressions such as 'in case'. The feature filters out cases like '*The question is whether we want to risk a nation of Misratas.*'

Expressive speech act verbs

As mentioned before, expressive speech act verbs have been partially included in the analysis, in the form of verbs that express blame or welcome. If someone condemns something, the utterer implies that something should not be done, and vice versa, if the speaker welcomes or encourages something, she expresses an implied agenda for action. Below, these verbs will be referred to as expressives.

Feature 11: if an expressive verb candidate is in the sentence. The list of verbs that express the speech act of blaming comprises 'blame', 'condemn', 'deplore', 'decry', 'denounce.' The list of verbs that express the meaning of encouraging includes 'encourage', 'welcome', 'applaud'. Other words with similar meaning have been excluded for the same reasons as commissive and directive speech act verbs. The list of excluded words comprises 'admonish', 'reprimand', 'endorse', 'countenance' and some others.

Feature 12: if an expressive verb candidate is tagged as a verb.

Feature 13: if an expressive verb is in auxiliary relation with the particle 'to'. This feature targets sentences such as '*They are to blame*' which do not express agenda for action.

Agenda for action communicated with temporal expressions

This group of features captures agendas expressed as '*The time to stop killing is now*' or '*It is the moment to go.*' The features are also supposed to filter out those propositions expressing duration, '*it takes long time to stop killing.*'

Feature 14: if the word 'time' or 'moment' is in the sentence.

Feature 15: if the word 'time' (or 'moment') is modified by a verb. This feature helps to distinguish between 'Time flies like an arrow' and 'The time to stop killing is now.'

Feature 16: if the word 'time' or 'moment' is a direct object of verbs 'take', 'require' or 'need.'

Grammar imperative and direct speech

The following group of features deals with sentences that express agenda for action explicitly as grammatical imperatives. In English, this is usually realized by putting the verb at the beginning of a sentence or a clause and omitting the subject. For instance, '*Fight them*' or '*He exclaimed: 'Fight them!*'' In order to disambiguate imperative and interrogative sentences (that also start with a verb), the features have been designed to check sentence-final punctuation. Relying purely on *n*-grams, would have most likely misclassified such sentences as there are no reliable lexical patterns or indicators.

Feature 17: if the sentence or the clause starts with a verb.

Feature 18: if the sentence finishes with a question mark.

Feature 19: if the sentence finishes with an exclamation mark.

Sentences with 'please'

These features target polite requests and favour seeking, such as '*Would you, please, open the window?*' They also distinguish between 'please' as the verb with the meaning 'to gladden or to rejoice' and as an interjection. '*They pleased us by coming over*' versus '*Please, come over.*'

Feature 20: if word 'please' is in the sentence.

Feature 21: if 'please' is tagged as an interjection.

Hint adjectives

The next group of features concerns adjectives that express the meaning of necessity or senselessness. They might be markers of agenda for action in sentences such as '*It is important to continue peace talks*' or '*The war is unacceptable.*' Seventeen adjectives have been included in the present work (see Appendix 2 for a complete list). The initial set of hint adjectives was defined based on the training corpus. Thereafter, it was enlarged based on relevant synonyms ('The Merriam Webster dictionary of synonyms and antonyms', 1992). Finally, adjectives that negatively affected classification results were excluded. For instance, 'unavoidable' is one of the synonyms of the word 'necessary', but unlike the latter, which is usually a strong indicator of an agenda for action, the former usually appears in sentences like '*He also tried to put this into the context of such attacks now being an unavoidable part of life in the world's biggest cities*', where no illocutionary force is present.

Feature 22: if a hint adjective is in a sentence.

Feature 23: if a hint adjective modifies a copula verb or is modified by the preposition 'as'. The feature targets sentences like '*They see the war as unacceptable*' or '*The war is unacceptable.*'

Agenda for action communicated as 'something won't stand anymore'
The following features are supposed to deal with sentences that express the idea that 'something cannot or should not hold anymore': *'They won't stand it anymore.'* The intuition is that, if someone says that they are tired of something and will not stand it anymore, the utterer will most likely do something to change the unfavourable situation. In natural language, in order to express this type of information, usually two components are required; a verb expressing an intention ('can' and 'go' for 'to be going to'), and the verb meaning 'stand' ('stand', 'withstand', 'endure', 'tolerate', 'put up'), e.g., *'I can't endure it anymore'* or *'They are not going to stand such a treatment.'* This group of features checks whether these verbs are present in a sentence and whether other relevant conditions are met.

Feature 24: if a verb expressing a speech act of planning has an open clausal complement expressed by a verb that means 'to stand'.

Feature 25-26: if the verbs meaning 'to plan' or 'to stand' is negated. This feature will be used for further investigation as to whether negation affects the motivational semantics of a sentence.

Feature 27: if a verb with the meaning of 'to stand' is in future tense. This feature checks whether the future tense can serve as an additional hint at the illocutionary force as in *'They won't stand it anymore.'*

Agenda for action with hint nouns
The following group of features captures agendas in such a form as *'Peace is the only solution'*. These structures implicitly call for actions to establish peace. The features are based on the list of hint nouns, which includes such words as 'answer', 'way', 'solution', etc. These features also disambiguate instances where hint nouns are actually verbs, as in, *'Peace is the answer'* versus *'He answered the phone.'*

Feature 28: if a hint noun candidate is in the sentence.

Feature 29: if a hint noun candidate is tagged as a noun.

Feature 30: if a hint noun is modified by the preposition 'as', or fulfils syntactic function of a direct object. The feature captures sentences like *'Peace is considered as an answer'* or *'Peace is an answer.'*

Agenda for action with hint adverbs

If something is vital, important, or essential, it is most likely that the utterer implies a necessary course of action. Feature 31 tackles such sentences. In order to filter instances when the adjective is in the sentence, but there is no call to action, this feature checks whether the adjective has a clausal complement, for example: *'Fruits are important for our health'* versus *'It will be vital that its terms are implemented.'* The list of adjectives that express the meaning of importance includes the following words; 'vital', 'important', 'significant', 'essential', 'substantial', 'principal', 'salient'. Such synonyms as 'critical' and 'indispensable' were excluded from the analysis for the same reasons as some speech act verbs.

Agenda for action expressed as 'To give a promise'

There are at least two ways in natural language to give a promise. First, with a verb (*'I promise to do so'*) or second, with a noun meaning 'promise' plus a verb expressing the meaning 'to give' (*'I give you a promise to do so'*). The former will be dealt with by the group of features that targets speech act verbs. The latter will be captured by Feature 32, which checks whether a noun 'promise' is a direct object of a verb 'to give'. The word lists cover different ways to express the meaning 'to give a promise.'

3.2.3 Data origin

As mentioned before, the data in the developed corpus originates from four different sources: traditional mainstream media, texts of strategic actors, parliamentary debates, and Twitter. All sources feature very specific and distinguishable language peculiarities. For example, consider that journalistic texts are expected to be written in standard literary language, while tweets, limited to 140 characters[13], are concise, full of abbreviations, typos and slang, are ignorant of grammar and orthographical rules. Political debates usually take the form of a polylogue which often results in ellipsis. UK parliamentary debates, for instance, feature many polite references such as 'my dear honorary

13 The limit was extended to 280 characters after the data collection and analysis had been performed.

friend' and interrogatives to express opinion: *'Has the Prime Minister made it clear to President Obama that in no way does this country support any attack that could come before the un inspectors have done their job?'* To capture these differences, the data origin has also been included as a feature in the classification routine. In a manner similar to linguistic features, data origin is a binary feature – a sentence either belongs to one of the four sources (feature value one) or not (feature value zero).

3.3 Stop words and lemmatization

As discussed in Section 2.1.3, stop-words removal and stemming / lemmatization can significantly improve the performance of a model, if the unit of analysis is larger than a sentence, or if a stop-words list was tailored to a specific corpus. Indeed, functional words such as particles or prepositions are irrelevant when identifying a topic or a genre of a document. However, they might be quite informative when distilling grammar patterns (Balakrishnan & Ethel, 2014; Silva & Ribeiro, 2003; Zaman et al., 2011). For sentence classification, especially for deciding whether it expresses an agenda for action, morphological and syntactic information is nearly as important as lexical patterns. Linguistic features described in Section 3.2.2 were designed to capture grammar and syntactic peculiarities and to emphasise their role in classification. However, linguist features cannot capture and address all of them. For instance, the verb 'push' can express an agenda for action as in *'to push the senators to vote for the intervention'*. However, due to its ambiguity, the verb 'push' was excluded from the list of hint speech act verbs: It is impossible to formulate a rule that will disambiguate the example above from *'to push the door to leave'* with reasonable precision. *'They pushed for intervention'*, if lemmatized will be normalized to *'They push for intervention'*, as a result 'push' as a verb in past tense will be regarded as a noun 'push' as in *'They felt a strong push when a bomb fell'*. It is very likely, therefore, that both sentences will be assigned the same label, which is wrong. Keeping the verb in past tense can give the classifier a hint that when 'push' is a verb (i.e., when there is 'pushed' in a sentence), the sentence is more likely to express an agenda for action. However, it still does not solve the problem of assigning correct label to the sentence *'to push the door to leave': It* will be most likely misclassified as an agenda for action. The same holds for stop-words (Khoo et al., 2006; Saif, Fernández, He, & Alani, 2014). *'To push for intervention'* is an agenda for action embedded in the phrasal verb 'push for'. After stop-words removal, preposition 'for' will be gone thus rendering disambiguation from *'to push the door'* impossible.

A number of experiments revealed the following optimal set-up. For the first step of classification, neither stop-words removal nor lemmatization or stemming have been performed. For the second step stop-words have been removed (Stalpouskaya & Baden, 2015). Moreover, the stop-words list has been tailored especially for the purpose of the current research; specifically, negators that traditionally are included into stop-words have been kept as meaningful words.

3.4 ML algorithms for agenda for action extraction

As explained in Section 2.1.4, there exists a handful of machine learning algorithms. The appropriateness of the selected algorithm depends greatly on the learning objective, nature, amount and complexity of data, and computational resources (Aggarwal & Zhai, 2012; Manning, Raghavan, & Schütze, 2009; Zhu, Vondrick, Fowlkes, & Ramanan, 2015). After examining a body of literature on text classification (c.f., Aggarwal & Zhai, 2012; Carvalho & Cohen, 2005; Colas & Brazdil, 2006; Hassan, Rafi, & Shaikh, 2011; Khoo et al., 2006; Pawar & Gawande, 2012; Qadir & Riloff, 2011; Ruiz & Srinivasan, 2002; Sebastiani, 2002) and running a number of experiments, the following algorithms have been chosen for the present thesis: *k*-Nearest Neighbours (*k*NN), Decision Tree (DT), Naïve Bayes (NB), Support Vector Machines (SVM), and Multilayer Perceptron (MLP), the last being an example of a deep learning algorithm. The scikit-learn (Pedregosa et al., 2011) implementation of all the classifiers with default parameters (unless mentioned otherwise) has been used. In the following sub-sections I will explain in detail the working principles of each algorithm.

3.4.1 Naïve Bayes (NB)

NB determines most likely class ownership using the following intuition. Firstly, it places a sentence to the class which has the highest probability (in the case of the first step of classification, we have two classes, and the non-agenda one has a higher probability). Then, the algorithm considers each feature and the probability of having said feature in a given class, thus increasing or decreasing the probability of assignment to each class. In the end, the class with the highest probability wins. It can be computed using the following equation (Equation 7):

$$y = \arg\max \left[\log P(y)\, P + \sum_{1 \le i \le N}^{N} \log P\,(x_i \backslash y) \right]$$

Equation 7. Class probability computation for Naïve Bayes

Where *y* is the class label (agenda for action or not), *P(y)* – a probability of class *y*, *i* – feature's number in a row, *N* – total number of features, $P(x_i|y)$ – a probability of having the *i*-th feature in class *y*

Figure 2 illustrates this logic (Bird, Klein, & Loper, 2009):

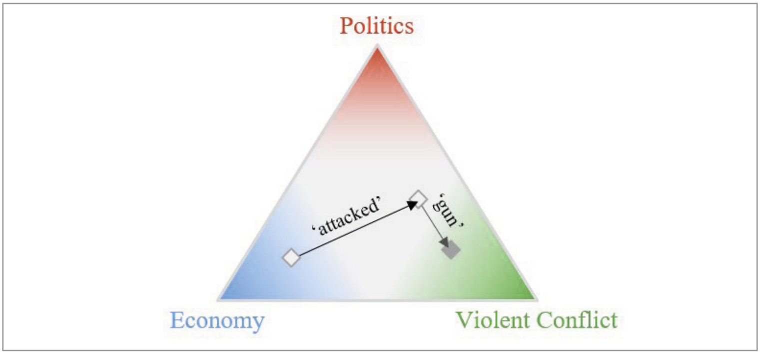


Figure 2: Naïve Bayes classification decision logic

The imaginary dataset consists of documents of three categories: 'Economy' (the majority class), 'Politics' and 'Violent Conflict'. First, the document is placed in the category 'Economy' as the most probable class. Then, the first feature comes into play which is the presence of the word 'attacked'. It makes the classifier fairly certain that the document does not belong to 'Economy', but two other classes – 'Politics' and 'Violent Conflict' are equally probable. Finally, the last feature is being considered – the presence of the word 'gun' which places the document to the class 'Violent Conflict', with great likelihood.

The weakness of NB is that it assumes the independence of each pair of features, which usually does not hold. Despite this limitation, NB has performed well in text classification tasks (Hassan et al., 2011; McCallum & Nigam, 1998; Rennie, Shih, Teevan, & Karger, 2003; Ting, Ip, & Tsang, 2011; Zhang, 2004).

3.4.2 *k*-Nearest Neighbours (*k*NN)

The *k*NN classifier decides which class assignment based on the class of the majority of its *k* nearest neighbours. *K* is a hyper-parameter defined experimentally. To identify the nearest neighbours, distance measurements are needed such as Manhattan, Euclidean, or Minkowski distance (Walters-Williams & Li, 2010). This intuition is reflected in Figure 3:

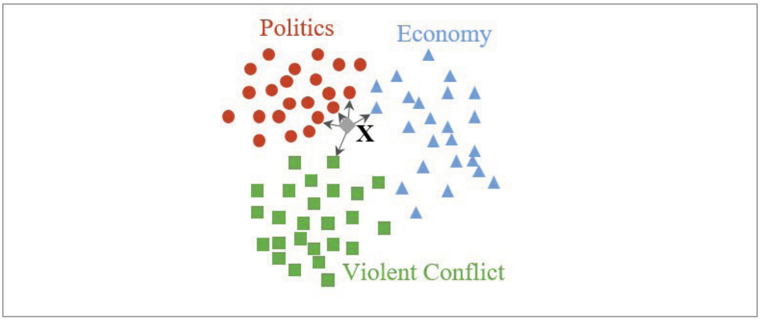


Figure 3. Visualisation of *k*NN algorithm with k=5

The dataset consists of three classes 'Politics', 'Economy' and 'Violent Conflict', in order to assign a class to a document X, one has to measure the distance to its 5 nearest neighbours. Among them, three belong to 'Politics' (red circles), one to 'Economy' (blue triangles) and one to 'Violent Conflict' (green squares), meaning that the document X is most probably a red circle and it will be assigned to 'Politics', with its nearest neighbours. The main weakness of this algorithm is that it is computationally time-consuming (Aggarwal & Zhai, 2012; Colas & Brazdil, 2006; Soucy & Mineau, 2001; Trstenjak, Mikac, & Donko, 2014).

3.4.3 Decision Tree (DT)

The decision making process of the DT classifier is the most similar to the human mental process. It can be explained using the dataset 'news' in which the decision what the news is about is made based on three determinants: domain, medium and means (Figure 4).

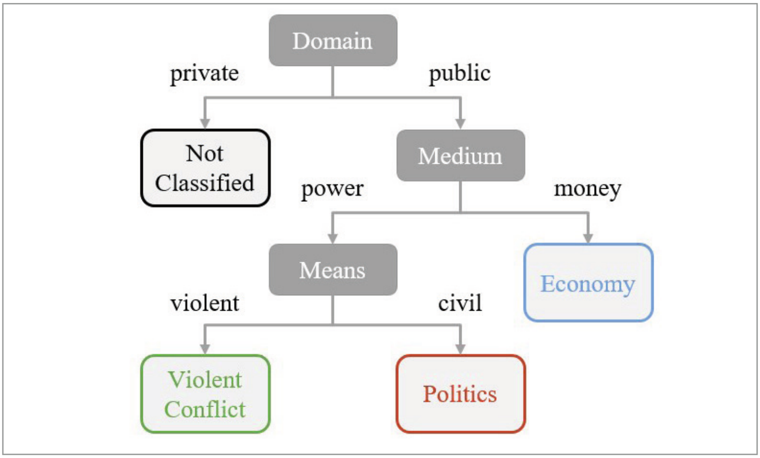


Figure 4: Visualisation of decision tree logic

Each node of the graph represents a feature considered during decision making. In this case there are three features: domain, medium and means. Each edge is the value that the feature may take (e.g., the feature domain can take values 'private' or 'public'). The aim of the model is to learn a tree that generates the shortest distance between top node and bottom leaf, i.e. the feature domain has been chosen as the top node, as, if its value is 'private', no other features need to be considered, while if medium, for instance, had been chosen as the top node, there would have been more steps required before reaching the bottom leaf (Hall et al., 2009).

The biggest disadvantage of DT is its instability. Even the slightest changes to data may result in a completely different tree, which therefore cannot be assumed to be the optimal one (Aggarwal & Zhai, 2012; Johnson, Oles, Zhang, & Goetz, 2002; Vateekul & Kubat, 2009).

3.4.4 Support Vector Machine (SVM)

SVM is also known as a large margin classifier. Its mechanism can be best explained with the help of an image (Figure 5):

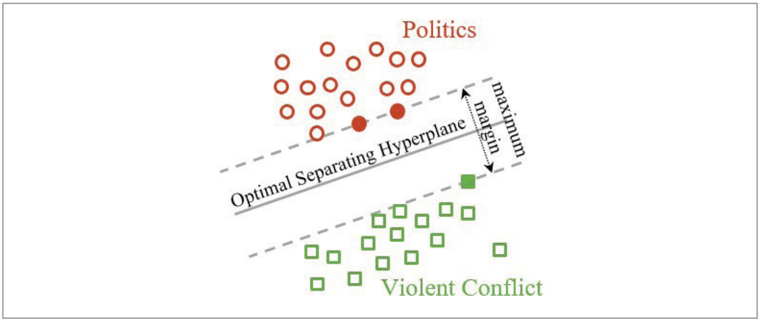


Figure 5: SVM ensures the largest margin between the decision hyperplane and data points

There exists a great number of ways to separate the two given datasets (red circles and green squares). Multiple hyperplanes can be drawn between them. SVM draws a dividing hyperplane in such a way that the distance between the plane and each data point is maximised. In other words, it creates a large margin between the hyperplane and each data point.

SVMs have proved to be one of the most powerful text classifiers. They can be fine-tuned to handle multiclass classification, non-linear and middle-sized datasets (Colas & Brazdil, 2006; Dumais, Platt, Heckerman, & Sahami, 1998; Hassan et al., 2011; Joachims, 2002; Maas et al., 2011; Manning et al., 2009).

3.4.5 Multilayer Perceptron (MLP)

MLP (Rosenblatt, 1958) is an example of a neural network-type deep learning algorithm. It consists of an input layer, one or more hidden layers and an output layer (see Figure 6). The input layer denotes the data units presented as feature vectors. Each node in the input layer corresponds to a feature vector which represents a data item, a sentence in our case. Each unit of a hidden layer is a neuron that applies some transformation rules to the values received from the previous layer and is initialised by an activating non-liner function. The output layer receives the input from the last hidden layer and transforms it into the result classification label (Manning & Schütze, 1999; Minsky & Papert, 1990; Ruiz & Srinivasan, 2002; Sebastiani, 2002).

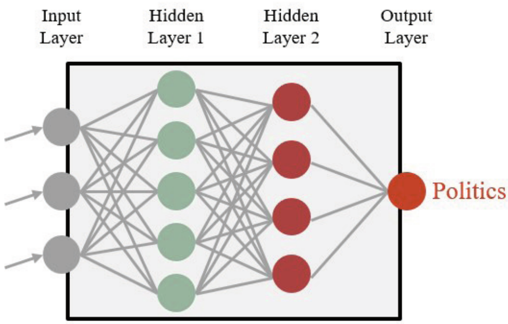


Figure 6: Multilayer Perceptron

Even though the MLP algorithm possesses a number of downsides, requiring fine-tuning of numerous parameters, is sensitive to feature scaling and needs much training data (Goodfellow et al., 2016), it has become the mainstream approach to text categorisation. For the present work MLP with two hidden layers five neurons in each and regularization term[14] alpha = 1e-5 was used.

14 Regularization term is any modification made to a learning algorithm that is intended to reduce its generalization error but not its training error (Goodfellow et al., 2016, p.117).

3.5 Classification results

In Table 4, I present classification scores for the described classifiers with different sets of features. In the first step, sentences are classified as agendas for action. Of all sentences in the corpus, 25% were used for testing, comprising 3,719 items, of these, 1,235 items were agenda for action and 2,484 items were not agendas for action. Grey cells denote the best scores. The scores for iterations in which only the data origin was used as a feature are omitted here, as they were extremely low.

	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
	NB (TF-IDF+LF+data origin)			SVM (TF-IDF+LF+data origin)			*k*NN (TF-IDF+LF+data origin)			DT (TF-IDF+LF+data origin)			MLP (TF-IDF+LF+data origin)		
agenda for action	0.94	0.33	0.49	0.74	0.68	0.71	0.63	0.58	0.6	0.64	0.7	0.67	0.73	0.71	0.72
not agenda for action	0.75	**0.99**	0.85	0.85	0.88	0.87	0.8	0.83	0.81	0.84	0.81	0.82	0.86	0.86	0.86
average	0.81	0.77	0.73	0.81	0.82	0.81	0.74	0.75	0.74	0.78	0.77	0.77	0.82	0.82	0.82
	NB (TF-IDF+LF)			SVM (TF-IDF+LF)			*k*NN (TF-IDF+LF)			DT (TF-IDF+LF)			MLP (TF-IDF+LF)		
agenda for action	**0.95**	0.29	0.44	0.76	0.72	**0.74**	0.57	0.39	0.46	0.6	0.7	0.65	0.68	**0.75**	0.72
not agenda for action	0.74	0.99	0.85	0.86	0.89	**0.88**	0.74	0.86	0.79	0.84	0.77	0.8	**0.87**	0.83	0.85
average	0.81	0.76	0.71	**0.83**	**0.83**	**0.83**	0.68	0.7	0.68	0.76	0.75	0.75	0.81	0.8	0.8
	NB (TF-IDF+data origin)			SVM (TF-IDF+data origin)			*k*NN (TF-IDF+data origin)			DT (TF-IDF+data origin)			MLP (TF-IDF+data origin)		
agenda for action	0.87	0.39	0.54	0.71	0.56	0.62	0.69	0.06	0.11	0.57	0.66	0.62	0	0	0
not agenda for action	0.76	0.97	0.85	0.8	0.89	0.84	0.68	0.99	0.8	0.82	0.76	0.79	0.67	**1**	0.8
average	0.8	0.78	0.75	0.77	0.78	0.77	0.68	0.68	0.57	0.74	0.72	0.73	0.45	0.67	0.53
	NB (TF-IDF)			SVM (TF-IDF)			*k*NN (TF-IDF)			DT (TF-IDF)			MLP (TF-IDF)		
agenda for action	0.87	0.41	0.55	0.74	0.68	0.71	0.73	0.03	0.05	0.6	0.61	0.6	0.78	0.48	0.6
not agenda for action	0.77	0.97	0.86	0.85	0.88	0.86	0.67	0.99	0.8	0.8	0.79	0.8	0.78	0.93	0.85
average	0.8	0.78	0.76	0.81	0.81	0.81	0.69	0.67	0.55	0.74	0.73	0.73	0.78	0.78	0.77
	NB (LF)			SVM (LF)			*k*NN (LF)			DT (LF)			MLP (LF)		
agenda for action	0.44	0.04	0.08	0.52	0.07	0.13	0.34	0.25	0.29	0.47	0.14	0.22	0.53	0.06	0.11
not agenda for action	0.67	0.97	0.79	0.68	0.97	0.8	0.67	0.77	0.72	0.68	0.92	0.78	0.68	0.97	0.8
average	0.59	0.66	0.56	0.62	0.67	0.57	0.56	0.59	0.57	0.58	0.53	0.5	0.63	0.67	0.57

Table 4: Precision, recall and f1 scores for DT, MLP, *k*NN, SVM and NB classifiers with different sets of features: TF-IDF only, TF-IDF and LF, LF only, TF-IDF with LF and data origin, TF-IDF with data origin

As reflected in Table 4, two setups have shown the poorest performance. Two classifiers – DT and *k*NN – possess the weakest prediction power for text classification tasks with any feature set. Likewise, using TF-IDF or linguistic features alone also reveals low performance scores.

NB has demonstrated the highest recall for not agendas for action with all three types of features and the highest precision for agenda for action with TF-IDF and linguistic features. However, NB performance was highly imbalanced, with a large gap between precision and recall scores for different categories, which resulted in a low f1 score.

SVM delivered the best results, outperforming *k*NN, DT and NB. MLP has also achieved high scores for the TF-IDF, TF-IDF combined with linguistic features and TF-IDF plus linguistic features and data origin feature sets, only just falling short of SVM. Potentially, with the availability of more training data, MLP could outperform SVM. Nevertheless, SVM with TF-IDF with linguistic features was the best performer. It yielded the best f1 score, which considers both precision and recall, measuring an overall balanced performance. Classification results support previous findings about the superiority of SVM in text classification among non-deep learning algorithms (Khoo et al., 2006).

3.6 The role of linguistic features

The use of TF-IDF scores has been one of the most common and well established approaches to text analysis in computational linguistics (Jones, 2004). As can be seen in Table 4, TF-IDF weighted n-grams and SVM as a classifier underperforms the best model by 0.02 – 0.03 in f1 score only. The addition that yields 0.02 – 0.03 points is the linguistic features described in Section 3.2.2. Their development is a tedious and time-consuming process, involving much data analysis and rule crafting. It also requires the use of additional NLP tools that costs additional computing time. In light of this, the question to be answered is whether additional linguistic features are beneficial to the model.

One of the criteria to estimate the power of a model is to evaluate it from the perspective of the bias – variance trade-off (Geman, Bienenstock, & Doursat, 1992; Manning et al., 2009). Some models suffer from under-fitting or high bias. If a model is very simple for the data at hand, it therefore assigns wrong labels to most data items. Such mislabelling can happen when, for instance, a linear model is used to separate non-linear data. Usually such models result in low training scores. On the other hand, some models suffer from overfitting or high variance. For instance, when a model is a polynomial function and attempts to correctly classify every data item, including noise and outliers. These models usually show high performance scores on the training set, but fall short in their ability to generalise and predict labels for new unseen data. Below are the examples of an under-fitting and overfitting models (Figure 7):

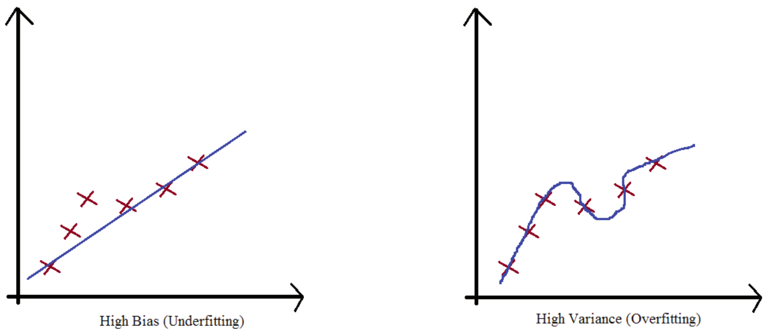
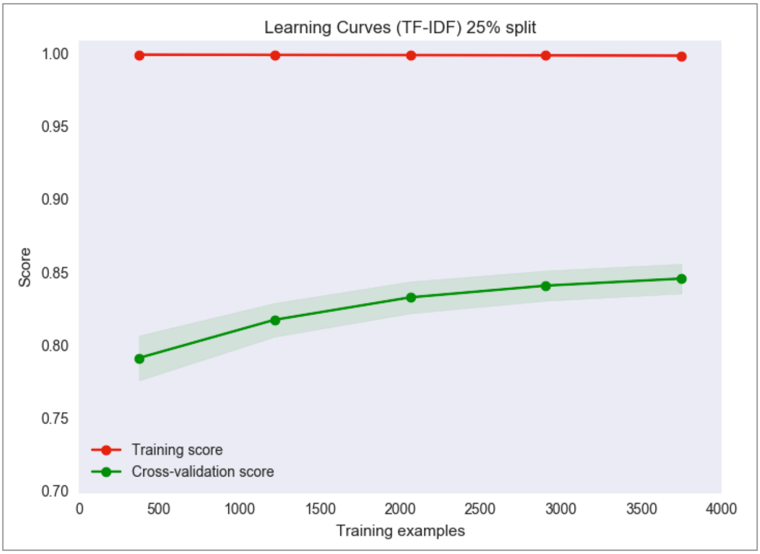


Figure 7: Visualisation of models that suffer from high bias (left) and high variance (right)

In order to create a powerful model that is able to correctly classify most items in the dataset and to be able to generalise classification to unseen data, one has to find the right trade-off between bias and variance.

Whether a given model suffers from high bias or high variance can be observed in learning curves. If one plots the change of classification score as the function of the size of the training and testing sets, the gap between the curves provides an estimate of bias, variance, or balance in the model. Figure 8 presents the learning curves for SVM classifier with only TF-IDF scores as features and the same classifier with TF-IDF scores enhanced with linguistic features and data origin.


Learning Curves (TF-IDF) 25% split
Score
1.00
0.95
0.90
0.85
0.80
0.75
0.70
Training score
Cross-validation score
0
500
1000
1500
2000
2500
3000
3500
4000
Training examples

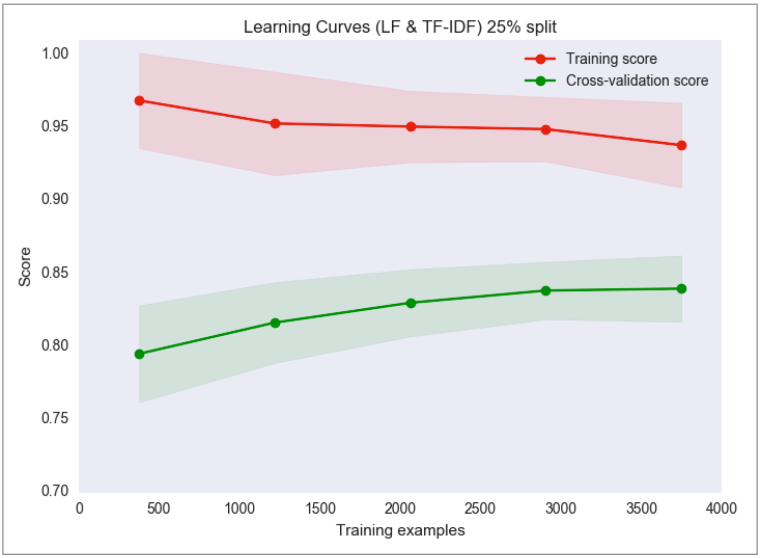

Learning Curves (LF & TF-IDF) 25% split
Score
1.00
0.95
0.90
0.85
0.80
0.75
0.70
Training score
Cross-validation score
0
500
1000
1500
2000
2500
3000
3500
4000
Training examples

Figure 8: Learning curves for SVM with different feature sets

The three graphs in Figure 8 present learning curves for SVM classifier with three different feature sets; TF-IDF only, TF-IDF combined with linguistic features, and TF-IDF with linguistic features and data origin. In the balanced model, the cross-validation curve is supposed to ascend with the growth of the test set and plateau at some point. The training curve is expected to descend and plateau around the same point as the cross-validation curve. There should exist a gap between the two curves, but it should not be too large (10 – 15 points on the y-axis). In the case of too wide a gap between the curves, the model overfits, it shows constantly high scores for the training data, supposedly also fitting noisy data, but fails to generalise to fit new data, leading to low scores on cross-validation set. The learning curves for the setup with only TF-IDF show such behaviour. In the case of two and three groups of features, the gap between the curves reduces. The training score decreases with the addition of new training examples, meaning that some data items are incorrectly classified, though such a finding is expected in the case of outliers or noisy data. This result means that the model is able to generalise well. This assumption is confirmed by the increased test score.

The learning curves reveal that the additional effort of implementing and deploying linguistic features does not only result in a slight performance boost, but also renders the model more powerful, overcoming the problem of high variance. For this reason, the SVM classifier with TF-IDF and linguistic features is employed for extracting agendas for action from text.

3.7 Classification error analysis

An analysis of the remaining misclassifications offers additional insights into the issues that could be addressed in future to improve the present approach. The SVM with TF-IDF weighted *n*-grams with $1 \leq n \leq 4$ and linguistic features has misclassified 577 sentences, 339 of which are agendas for action and 238 are not. The distribution of misclassified sentences per data category is as follows:

- strategic communication – 167
- social media – 150
- political communication – 122
- journalistic texts – 138.

The numbers reveal that no particular discourse style is more challenging for the model to classify correctly, as the misclassified sentences of this kind are spread evenly across discourses.

The classification errors can be split into several large categories.

Imperative sentences in social media

There are distinguishing features of the imperative mood. The first feature is the omission of the subject, which places the verb at the beginning of the sentence or clause. Second, some imperative sentences finish with an exclamation mark, which helps to disambiguate them from interrogative statements. For social media it is common to start a tweet with a hashtag, identifying the addressee of the utterance. In many cases, these sentences were misclassified most probably because the addressee was treated as the subject of the sentence, hence imperative mood was not recognised. ‘*‘#Congress: stop the political games!!*’, ‘*@Ashton5sos @xoxoLaurmani dont reply to a 5h stan were at war x1.*’ This problem could potentially be fixed by adding a pre-processing step and removing the first word if it starts with a # or @.

Complex grammar patterns
Some complex grammar structures, such as conjunctive mood, for example '*I think it would have been better if there were more concrete and clear messages*', were not recognised as expressing agendas for action. This problem could have been solved through the implementation of additional linguistic features that recognise peculiarities of this mood.

Implicit agendas for action
Due to ambiguity and the lack of formal hints, some sentences were misclassified. For example, '*The place to resolve the differences between the parties is through direct negotiation, not unilateral actions by either side*' was not identified as an agenda for action. Here, if the utterer had used the word 'time' instead of 'place', the sentence would likely have been classified correctly, as the word 'time' is an identified hint noun. In contrast, adding the word 'place' to the hint noun list would have produced even more misclassifications than correct results.

The majority of misclassified sentences comes from political communication texts and include implicit indirect agendas for action expressed as questions, such as '*Will my right hon. Friend impress on the authorities that have custody of the bodies that it is a matter not just of dignity, but of identification?*' Grammatically, this is an interrogative sentence which is not a typical way to express agendas for action. It originates from parliamentary debates and for texts of this type it is common to express agendas for action indirectly (Bavelas, Black, Chovil, & Mullett, 1990; Obeng, 1997). Moreover it is a sentence from a debate in the UK House of Commons, which features a more classical, polite, high-level English, in which register it is common to express requests as questions (c.f., Brown & Levinson, 2011[1987]; Harris, 2001; Murphy, 2014; Wilson, 1990). The feature 'data origin' was supposed to help the classifier to handle such examples, but in this case it has failed.

Incomplete lists of hint words

Some errors were caused by the incomplete lists of hint words. The verb 'appeal', for example, was not on the 'hint verbs' list, even though it appears in the agenda for action '*In 2014, UNHCR and partners appealed for US$210 million under the Central African Republic Regional Refugee Response Plan*' that was misclassified. Updating and enriching the lists of hint words is an iterative process, which requires constant testing and refinement. The lists may also vary given the volume and the nature of the data.

Wrong annotation

Some of the reported misclassifications were due to imprecisions and discrepancies in annotation. Some sentences were actually not misclassifications, but correctly recognised agendas for action. Consider the sentence '*Bloomberg: Debt crisis shows Puerto Rico needs to be a State or a nation.*', for example. The sentence expresses an agenda for action but was labelled as not agenda for action in the corpus. Even though wrong annotations affect the performance score, the fact that such sentences are still assigned correct labels shows the power of the model, as it has managed to avoid overfitting and is able to generalise well. This is also an example of the superiority of machine predictions to human decisions.

3.8 Assigning categories to agendas for action

3.8.1 Variety of taxonomies

While it might be valuable to know that agendas for action are present in a text or that the number of advanced agendas for action over a certain time span is bigger than usual, it is of great importance to know the nature of the actions that are called for. In other words, it is not only necessary to be able to extract them from texts, but also to classify them.

There exists no universal agenda for action taxonomy, it has to be tailored depending on the domain and discourse. If the object of the analysis is a corpus of emails, then one possible agenda for action classification may be to classify them as reminders, requests, or plans. If agendas for action are extracted from medical texts, then they may fall into the classes of pharmaceutical prescription, referrals and future check-ups, or nutrition and physical activity recommendations. When analysing journalistic discourse on economy, it might be interesting to know whether there are calls for reform, tax increases, or support of SME. In principle, there exists an infinite number of ways to classify agendas for action. In this chapter I propose one possible taxonomy for agendas for action in texts about war and violent conflict, as well as demonstrate its automatic extraction.

3.8.2 Agenda for action classification in texts about war

In present thesis I deal with texts about war and violent conflict stemming from four different discourses; politicians and strategic communicators calling for escalation or de-escalation, journalists charting different courses of possible action, and social media audiences rally toward hostility or cry out for calm and peace. It is of particular value and interest to analyse agendas for action advanced by different parties or actors involved in the conflict directly or indirectly, as knowing the nature of action or inaction being called for may help to track the dynamics of the conflict, perhaps even predicting phases of escalation and de-escalation.

The following taxonomy has been developed using inductive and deductive approaches. First, a number of classes based on the study of war and peace journalism (Galtung, 2013; Goretti, 2007; Hanitzsch, 2004; Lee & Maslog, 2005) have been drawn (termed agendas for peace and agendas for war). Further fine-tuning of the taxonomy was data-driven. Some new categories introduced (such as agendas for no n-action or general agendas for action), and others were merged together in the case of very few instances in the corpus, or if they were highly ambiguous. Finally, the following classification was developed and the corpus was annotated. The classification system includes nine classes:

- Peaceful solution and de-escalation – this class includes agendas for peace, ceasefire, and calls to stop fighting or violence. For example, '*The United States is closely monitoring developments in the Kyrgyz republic and calls for a rapid restoration of peace and public order, State Department spokesman Philip Crowley said in a statement Saturday.*' The utterer emphasizes that violence is not acceptable, thus hinting that the only appropriate solution is a peaceful one. Further in the text this class of agendas is referred to as 'agendas for de-escalation'.
- Violent solution and escalation – this category is opposite to the previous one. It includes agendas for military action, violence, killing, escalation, and physical destruction. An example is found in the sentence '*Yemen foreign minister calls for Gulf Arab military intervention.*' In this example, an agenda to destroy the state of Israel is advanced, which is a blatant call to a violent action. Agendas of this class can address individuals and groups, both agendas to kill a specific person as well as destroy a group of people (e.g., a state, an army unit). Agendas of this class are referred as 'agendas for escalation'.
- Involvement, dialogue, support and help – this class is semantically close to the first category, in the sense that these agendas call for something good and positive, though addressing physical, economical or humanitarian needs, such as cooperation, negotiations, food supplies, medical help, prisoners exchange, and financial support. An example from the corpus is '*Eradication of poverty should be the main priority of humanitarian action.*' Agendas of this class are termed 'agendas for help and support' in this dissertation.
- Punishment, sanctions and toughness – agendas from this class are the opposite of the previous category. They call for a tough stance and (non-violent) coercion including legal restrictions and punishment. Consider, for example, '*Joins the Foreign Affairs Council of 22 June 2015 in calling on the Vice-President of the Commission/High Representative of the Union for Foreign Affairs… to prepare a list of targeted restrictive measures and visa and travel bans against those responsible for acts of violence, repression and serious human rights violations...*' Agendas of this class are referred to as 'agendas for punishment'.

- Ignorance, exclusion, self-defense and protest – the agendas in this category are semantically close to agendas for escalation and agendas for punishment as they do not call for help, cooperation or peace. Instead, these agendas call for ignorance of the other side by excluding, not attending, or by closing up and fending off. One such example is, *'The amendment should restrict military courts to trial of hardcore terrorists only, he added.'* Agendas of this class are termed 'agendas for ignorance'.

  Semantically, agendas for escalation, punishment and ignorance can be labelled as ones calling for negative treatment of an opponent. These three classes may comprise a scale of negative actions towards an opponent. Agendas for escalation call for physical destruction, agendas for punishment call for certain restrictions and actions to be taken towards the opponent, while agendas for ignorance rather call for an intended inaction and hostile attitude towards the opponent.
- General agendas for action and rhetorical questions – some agendas simply express dissatisfaction with the status quo and the intention to undertake action, though they do not specify a precise course of action. For instance, *'Now is the time to take action.'* Agendas of this class will be referred as 'general agendas'.
- Agendas for not doing – this class includes calls to avoid a certain action, as in *"We must not lower our guard, at any time', Prime Minister Manuel Valls told Parliament, adding that 'serious and very high risks remain'.'* Sentences criticising others for doing something also belong to this class; *'We condemn these barbaric crimes.'* In most cases there exists a formal check whether an utterance belongs to this category. If it can be rephrased in the structure 'modal verb + negator' and the meaning remains the same, then the proposition calls for not doing, as in the following example: *'The Department of State warns U.S. citizens of the risks of travel to eastern Ukraine'* can be transformed into *'The Department of State suggests that U.S. citizens should not travel to eastern Ukraine.'*
- Multiclass – complex sentences wherein each clause expresses a different agenda fall under this category. One example is *'The international community must break that habit, accept the Palestinian membership application, guarantee Palestinians a war crimes case, prioritize peace*

and end Israel's impunity – or see international law perverted further in ways that is certain to harm the entire world.' In this sentence, five clauses express agendas for action: 'Break the habit' is an agenda for not doing, 'accept membership' belongs to involvement and dialogue, 'guarantee a war crime case' falls under punishment, 'prioritize peace' is a call for de-escalation and 'end impunity' is also punishment. Noteworthy is that sentences that consist of multiple clauses, but express agendas from the same class (for example when all clauses call for de-escalation), are treated as a single agenda for action belonging to the respective class.

- Other – sentences that contain an agenda for action but are semantically ambiguous and are hard to classify into one of the above categories. For instance, *'The militants who massacred schoolchildren, beheaded soldiers and attacked defense installations have surely committed war crimes and must be dealt with as such.'* In this example, the exact meaning of 'must be dealt with as such' depends on the utterer and the extra-linguistic context: It might be a call to kill them (agenda for escalation), to prosecute (agenda for punishment) or two of them together (multiclass).

The categories of agenda for action described above contribute differently to content analysis and action prediction. While the first six categories can indeed forecast the outcome of a situation, the last three categories may not, though have been introduced for the sake of an automated classification routine. Agendas for not doing and multiclass agendas do not tell of an expected outcome or programmed action, but rather serve as a signal of uncertainty in text, as well as of a need for an in-depth qualitative analysis. The 'other' class is introduced so as to capture those instances that are clearly agendas for action, though the classifier could not confidently assign another class to them. The reason for difficulty with classification may be a new or complex way of expressing an agenda for action, which is not covered in the training corpus. It may also arise from erroneous syntactic analysis of a sentence. Chapter 5 of this thesis discusses this limitation and outlines ways to improve the current approach.

3.8.3 Distribution of agenda for action categories in the corpus

Table 5 presents the distribution of agenda for action categories in the developed corpus according to the taxonomy introduced in Section 3.8.2:

	Traditional media	**Strategic communication**	**Political communication**	**Social media**	**Total per category:**
Agendas for de-escalation	128	16	58	40	242
Agendas for escalation	99	1	3	57	160
Agendas for help and support	165	475	497	59	1196
Agendas for punishment	100	9	37	13	159
Agendas for ignorance	38	9	12	13	72
General agendas	75	27	32	16	150
Multiclass	43	65	1136	14	1258
Agendas for not doing	115	74	88	119	396
Other	157	177	705	245	1284
Total per discourse:	920	853	2568	576	4917

Table 5: Different categories of agenda for action per discourse in corpus

Inter-coder agreement for fine-grained classification has achieved the score of 0.45, regarded as adequate for computational linguistic studies (Landis & Koch, 1977; Marion, 2004). The confusion matrix is presented in Appendix 1.

The data are strongly skewed with some categories being underrepresented (such as agendas for ignorance) and others over-represented (other). In order to overcome this problem, hierarchical classification was performed. First, agendas for action were classified as promoting cooperative treatment (agendas for de-escalation and help and support fall under this category), restrictive treatment (agendas for escalation, punishment and ignorance) or other (all other categories merged). Then, sentences in the first pool were classified as calling for de-escalation or help and support, while sentences from the second pool were categorized as calling for escalation or restriction, and sentences from the third pool were classified as agendas for not doing and other. Thereafter, the sentences from the restriction group were further classified as agendas for punishment or ignorance, or multiclass. The rest were distilled from the 'other' group. Finally, propositions in the 'other' were classified into general agendas and other. The closeness of some categories is also highlighted by the confusion matrix (Appendix 1), which shows that sentences from categories such as 'punishment' have usually been confused with the category 'ignorance', and 'general' with 'other'. Inter-coder agreement score for four labels (cooperative, restrictive, other, not agenda for action) reached 0.513. For five labels (cooperative, restrictive, agendas for not doing, other and not agenda for action), inter-coder agreement reached 0.5.[15] Classification steps are shown in Figure 9:

15 As the coders have not been presented with merged categories and have been annotating the sentences with all ten categories (not agenda for action and nine categories of agenda for action), to have the same testset for agreement measurements, the scores for all the categories (including not agenda for action) are reported here.

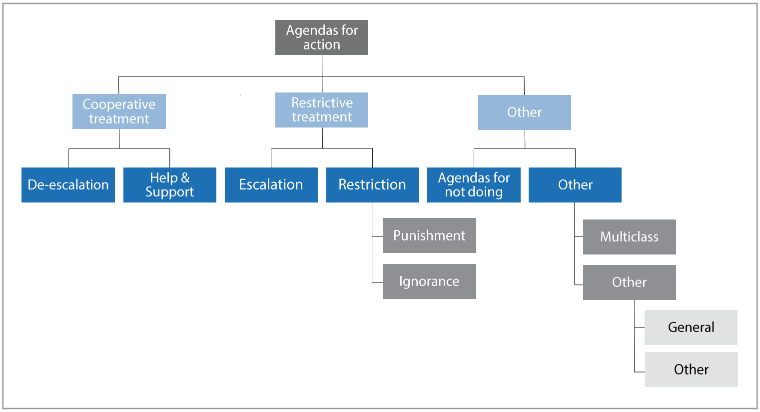


Figure 9: Steps in hierarchical classification of agendas for action

3.8.4 Automated categories assignment using ML algorithms

For the second classification step, the same classifiers as in the first step were used, namely, Naïve Bayes (NB), Decision Tree (DT), *k*-Nearest Neighbours (*k*NN), Support Vector Machines (SVM) and Multilayer Perceptron (MLP). However, there are two differences in the setup, compared to the first classification round. First, stop-words have been removed for the fine-grained classification. The stop-words list was tailored specifically for this task. For example, negators, which are usually included in such lists, as they do not contribute to semantic disambiguation, were considered meaningful in the current setup. Thus, they were excluded from the stop-words list as they are very strong indicators of agendas for not doing and can also disambiguate such sentences as *'The sanctions against the aggressor should be toughened'* and *'The sanctions against the aggressor should not be toughened.'* Second, given the scanty data, only 20% of the total data was used for testing. Third, as linguistic features are only helpful for discerning agendas for action from not agendas for action, they were not used in this step of classification. Classification results for three-label classification (cooperative treatment, restrictive treatment, other) are shown in Table 6:

	TF-IDF for NB			TF-IDF for SVM			TF-IDF for *k*NN			TF-IDF for DT			TF-IDF for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
Cooperative treatment	0.5	0.01	0.03	0.61	0.4	0.48	0.49	0.37	0.42	0.41	**0.51**	0.46	0.72	0.23	0.34
Restrictive treatment	**1**	0.02	0.05	0.65	0.3	0.41	0.38	0.26	0.31	0.47	**0.43**	**0.45**	0.92	0.14	0.24
Other	0.65	**0.99**	0.78	0.73	0.89	**0.8**	0.72	0.82	0.77	0.74	0.67	0.7	0.69	0.96	0.8
Average	0.63	0.64	0.52	0.69	**0.71**	0.68	0.63	0.65	0.63	0.63	0.61	0.61	0.71	0.69	0.63
	TF-IDF + data origin for NB			TF-IDF + data origin for SVM			TF-IDF + data origin for *k*NN			TF-IDF + data origin for DT			TF-IDF + data origin for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**Precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
Cooperative treatment	0.5	0	0.01	0.59	0.44	**0.51**	0.54	0.41	0.47	0.38	0.47	0.42	0.58	0.42	0.49
Restrictive treatment	0	0	0	0.65	0.32	0.43	0.46	0.37	0.41	0.44	**0.43**	0.43	0.89	0.1	0.18
Other	0.64	1	0.78	0.74	0.87	**0.8**	**0.75**	0.84	0.79	0.71	0.64	0.68	0.71	0.88	0.78
Average	0.55	0.64	0.51	0.69	0.71	0.69	0.67	0.69	0.67	0.6	0.58	0.59	0.69	0.69	0.65

Table 6: Precision, recall and f1 scores for fine-grained classification with three classes

Different models with different feature sets were more successful in different classification steps. For distinguishing between the categories 'de-escalation' and 'help and support', SVM with TF-IDF scores as features was used. To disambiguate between 'escalation' and 'restriction' MLP with TF-IDF and data origin as features was used. For classification of 'punishment' versus 'ignorance' *k*NN with TF-IDF scores performed best. Agendas for not doing were distilled using *k*NN with two feature sets. For the two last steps of the fine-grained classification (i.e. complex agendas with multiple calls comprised into one sentence, 'general' and 'other') SVM with two sets of features (TF-IDF + data origin) was best suited for the task.

Classification rates for hierarchical classification described above in comparison to all categories being assigned at once (nine-labelled classification) are presented in Table 7:

	Hierarchical			**All categories at once**		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**
Agendas for de-escalation	0.72	0.67	0.69	0.55	0.51	0.53
Agendas for help and support	0.93	0.95	0.94	0.61	0.5	0.55
Agendas for escalation	0.86	0.18	0.3	0.59	0.36	0.44
Agendas for punishment	0.78	0.97	0.86	0.53	0.29	0.38
Agendas for ignorance	0.8	0.31	0.44	0	0	0
General agendas	0.57	0.19	0.29	0.07	0.04	0.05
Multiclass	0.66	0.83	0.74	0.5	0.81	0.62
Agendas for not doing	0.8	0.09	0.16	0.5	0.41	0.45
Other	0.94	0.99	0.96	0.5	0.39	0.44

Table 7: Comparison of classification scores performed hierarchically and at once

As can be seen in Table 7, introducing hierarchical classification improved precision scores for all categories. Recall also improved for all but two, namely 'escalation' and 'agendas for not doing'. With fewer labels to choose from, it is statistically more probable that an algorithm will choose the correct label, which is seen in superior precision scores for hierarchical classification. The trade-off here is that by reducing the number of labels to be learned, the algorithm does not learn finer patterns and data peculiarities that pay into telling finer categories apart. This results in lower recall scores compared to precision. It means that when the algorithm classifies a sentence as one of the agendas, it is fairly confident in its classification decision, but picking all the instances of a given agenda for action from the dataset is hard for the algorithm.

For two categories, namely agendas for escalation and agendas for not doing, recall scores are better in one-step classification (0.36 and 0.41 respectively against 0.18 and 0.09). These two categories turn out to be the noisiest and most ambiguous ones from the point of view of the algorithm. Agendas for escalation are also poorly presented in the corpus (160 data points), they are observed in traditional and social media mainly, and are barely seen in strategic and political texts. All these makes it hard for the algorithm to learn to capture as many instances of them as possible, and providing the algorithm with less data does not allow it to learn all the peculiarities and patterns necessary to disambiguate agendas for escalation. Similarly, agendas for not doing are very diverse in themselves: As was mentioned in Section 3.8.2, there are at least two ways to express agendas for not doing – by using negators such as 'not' or by using speech act verbs such as 'blame'. For this category, reducing amount of training data by hierarchical classification did not allow the algorithm to learn these patterns, which has low recall as a result. Very low recall scores in case of agendas for escalation and agendas for not doing lead to lower f1 score in hierarchical classification compared to one-step classification.

Additionally, the algorithm is optimised against the f1 score, i.e., such a set of parameters is chosen that the highest f1 score is obtained. It means that there might exist such a parameter set, where precision or recall might be higher than presented in Table 7. In other words, there might exist a set up where recall is better for all categories in hierarchical classification, but that set-up would not have yielded the highest f1 score.

The scores for all the classifiers for all steps are shown in Appendix 3.

3.9 Conclusion

In this chapter, I have demonstrated the automated extraction of agendas for action from texts using machine learning approaches. This process involved many steps. First of all, the representative corpus had to be collected and annotated. The annotation process was more challenging than expected, as many cases from the corpus are highly ambiguous, resulting in a rather low inter-coder agreement score. Nevertheless, the developed corpus is the first attempt to collect and classify not only explicit agendas for actions, but also implicit ones.

The task of agenda for action extraction and classification was accomplished using machine learning. Although machine learning has been the mainstream approach in most scientific fields dealing with large volumes of textual data, communication science has mostly used traditional content analysis methods or lexicon based automatic approaches. The statistical model developed was rendered more powerful due to the linguistic features designed especially for the purpose of the present study. These linguistic features identify grammatical and lexical patterns in agendas for action and up-weight them among other features used in the model.

Agenda for action classification has proved to be a non-trivial task. First, the taxonomy taking into account the purpose of the task had to be developed. This included a definition of the classes of agendas for action that might be interesting for the purpose of this research. In this chapter, one of the possible taxonomies for agendas for action extracted from the coverage of violent conflict was introduced. The given taxonomy can be further adapted and expanded, especially with growth of the corpus. One of the biggest challenges is to classify agendas for action in an automated fashion. The developed corpus is still relatively small, with some categories being underrepresented. For this reason, classification rates may be fairly low. In order to boost performance, hierarchical classification was employed. Categories that are semantically close were first merged into hypernym buckets that were extracted primarily. Then, further classification steps were performed within each bucket in order to distil more fine-grained categories of agendas for action. In total, the fine-grained classification was completed in five steps which improved the performance of the model. For each of those steps, different algorithms with different feature sets revealed dissimilar performance, resulting in the final system consisting of multiple models.

To the best of my knowledge, the present study is the first to extract and classify highly complex contents, such as agendas for action. Given its novelty, there is significant room for improvement. To address these issues, the corpus needs to be cleaned and enlarged. It would also be desirable to have an inter-coder agreement in the range of 0.7 – 0.8. Also, at least 5,000 examples of agendas for action of each category would bolster precision of extraction and assignment. Secondly, the model can be further tweaked with the deployment of different algorithms (including neural networks). Linguistic features could also be improved, as more data might inspire the development of new features or may prove the redundancy of existing ones. Additionally, the hint word lists should also be constantly updated based on the results of data analysis. There is also a lot of potential in the application of word embeddings for classification of agendas for action (Mikolov, Chen, Corrado, & Dean, 2013).

In the following chapter I demonstrate, with concrete examples, the capabilities of the developed tool and the manner in which extracted information may be used to study texts regarding conflict.

4 Interpretation of agendas for action in news coverage of chemical weapons crisis in Syria in 2013

In this chapter I demonstrate the application of the principles and methods outlined in previous chapters to content analysis. I highlight that agendas for action extracted from news coverage may be used to better understand and predict conflict dynamics, as well as to observe foreign policies in real time. I use the 2013 chemical weapon attacks in Syria as a case study.[16] The situation in Syria at this time was very unclear and contradictory. Multiple, frequently contesting agendas were advanced by different parties, each trying to understand the events and hash out a path forward with Syria (Baden & Stalpouskaya, 2015b). The case can properly demonstrate the full potential of analysing agenda for action. All examples of agenda for action analysis, as well as conclusions drawn, should be treated as guidelines for the nuanced and multiplex nature of agenda for action analysis, not as ultimate findings or facts. Also, the ways in which agenda for action can be used and interpreted should not be limited by the examples provided in this chapter.

4.1 Timeline of events in Syria[17]

4.1.1 Syria in 2011–2018

The civil war in Syria started as civil uprising against the government of Bashar al-Assad, with demands to release political prisoners, of democratic reforms and increase of freedom. Steadily, slogans turned less peaceful, even calling to overthrow Assad's regime. Protests also spread beyond the capital city of Damascus and, by April 2011, had taken

16 This study has also been a part of the FP7-EU Project infocore (www.infocore.eu, Grant Agreement No. 613308). The aim of the project was to investigate the role of mass media in conflict areas on the example of six war or post-war regions: Burundi, Democratic Republic of Congo, Kosovo, Macedonia, Israel-Palestine and Syria.

17 This timeline draws upon following sources: Fröhlich (2018), Balanche (2018), "A Timeline of the Syrian Conflict as It Enters Its Eighth Year," (2018) place in about twenty Syrian cities. The government has suppressed unrest rather violently, causing multiple causalities among civilians (Holliday, 2011).

Civil unrest was followed first by military insurgencies in summer 2011, when several of the defected generals formed the Free Syrian Army, marking the establishment of oppositional military forces. Since then, clashes between the government and the opposition have become increasingly violent. The situation in Syria at this time had already been referred to as civil war by certain media outlets (Kennedy & Jordans, 2011). However, the first official statement by UN officials proclaiming that Syria was in civil war was released in June 2012 ('Syria in civil war, says UN official Herve Ladsous', 2012).

The situation became even more complex in 2014, when Islamist groups including Islamic State of Iraq and the Levant (ISIL) joined the warfare as a third party, opposing both the government and the opposition forces. At approximately the same time, the presence and participation of the international community became even more prominent. The US claimed to have supported the rebels by providing military training, while Russian, Turkish and Iranian armies were fighting on the side of the government. These same countries backed the ceasefire agreement, which was distracted by the first direct air attack by the US, in April 2017. By the end of 2017, another ceasefire attempt between the opposition and official forces was made, and the main forces of the Assad army, together with foreign militia, fought against ISIL.

2018 started with advances of Turkish army aiming to free the southern part of the country from Kurdish rebels. In April, another chemical attack was reported. Similarly to 2013 incidents, the wrongdoer was not identified. In the following months, military activities were carried out by various parties. Russian, American, Israeli, Turkish, Syrian armies as well as Kurdish-led and rebel-led forces were trying to gain momentum and to secure their positions.

4.1.2 Chemical weapon attacks in 2013

Arguably, the most noticeable event in foreign media during Syrian civil war in 2013 was the chemical weapon attacks near Aleppo on the 19th of March, and in Ghouta on August 21st. These attacks fuelled international debate, causing different parties to call for varied solutions to the crisis, rendering the agenda landscape extremely heterogeneous. The use of the nerve toxin, sarin gas, in both attacks was confirmed by the reports of Russian experts (Louis Charbonneau, 2013), as well as the UN investigators (UNHRC, 2014). However, the mastermind behind these incidents remains unclear. While the government and the rebels blamed one another, the report by Russian investigators held the opposition accountable. The UN report restrained from drawing conclusions in this regard.

The Syrian storyline dominating foreign media outlets in 2013 looked as follows. The first rumours about conventional weapons having been used in Aleppo were circulated in March, however they were neither confirmed nor denied. The international community speculated on this topic, and even considering military intervention. Russian experts collected samples from the attacked territory and conducted analysis in an OPCW certified laboratory. They concluded that sarin had been used and that the attack had been carried out by the Syrian opposition. The Russian ambassador to the UN shared insights from the investigation at the UN press conference on the 9th of July. Around the same time, the Syrian government requested that the UN send investigators to Syria to examine the site of the alleged chemical weapon attack. In the meantime, the US and UK, as key international players, contemplated military intervention in Syria to deter the usage of conventional weapons. On the 21st of August a much larger-scale chemical weapon attack took place in Ghouta. At this point, the intervention of the international coalition seemed inevitable, however, it did not take place. Although then UK Prime Minister James Cameron was in favour of an airstrike, he failed to convince the government to do so, resulting in the House of Commons voting against the intervention. The vote took place on the 29th of August. The final step that prevented further escalation was the diplomatic efforts of the Russian Federation, a Syrian ally. It convinced Syria to surrender their chemical weapon stockpiles and to join the Chemical Weapons Convention on the 14th of September. The Syrian disarmament was the main reason the intervention did not take place. Shortly after, the UN published its report confirming the use of sarin in Ghouta. A follow-up report by the UNHRC was released in February 2014, stating that the chemical used in Aleppo attacks was also sarin.

The timeline of the key events in 2013 with respect to the chemical weapon crisis is shown below in Figure 10:

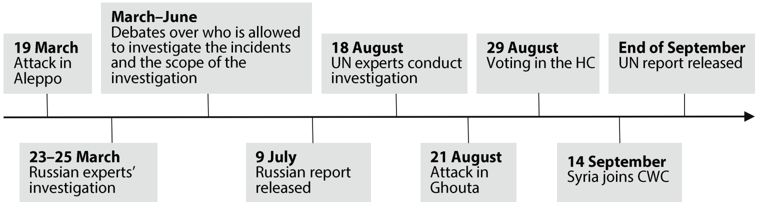


Figure 10: Main events in Syria in 2013 wrt Chemical Weapon attacks

4.2 Agendas for action in the chemical weapons crisis news coverage

The situation in Syria in 2013 was very unclear and contradictory. There were multiple opinions and suggestions regarding a) whether chemical weapons had been used; b) who was responsible for the attacks; and c) how to handle the situation on an international level. Several years after the incidents, some of the questions have been answered, namely, it has been proven that chemical weapons had been used, and it is known that no military intervention happened during or right after the crisis. The party responsible for the attacks, however, remains unknown.

When tracking the events of conflict, it is even more challenging to make sense of multiple sources of information. We do not usually have access to the locations of the events, but are instead provided with this information by media. Different media outlets tend to present information from different angles, they also represent different opinions and policies, depending on the country of their origin and any political affiliation. Based on the framing of the issue by the outlet, different courses of action can follow.

As has been shown in previous chapters, identifying agendas for action in public discourse is crucial to foresee the outcome of a course of action. In the case of the Syrian chemical weapon crisis, extracting agendas for action from news texts can help to better understand the events on the ground, as well as to build a more informed assumption as to the resolution of the issue.

4.3 Agendas for action in *The New York Times* and *The Guardian* during Syrian chemical weapons crisis

In order to test the methodology developed in Chapter 3 of the present thesis, I have used the coverage of the Syrian chemical weapons crisis in the American media outlet *The New York Times* (NYT) and the British outlet *The Guardian*. The model that was trained and described above has been applied to these two newspapers to showcase the deployment of automated extraction of agendas for action and including analysis and explanation of the situation on the ground. This chapter also offers other uses for the developed methodology. The reason for selecting *The NYT* and *The Guardian* for the analysis was the fact that both the USA and the UK were key players in the crises. Both outlets are examples of mainstream elite papers that closely follow foreign policy debates and, presumably, report on the events in an objective and largely unbiased manner. Another reason is a rather pragmatic one. While there exist ample computational tools and resources to process English language data, there is still a lacuna regarding other languages. The analysis presented in this chapter could benefit greatly from inclusion of agendas for action extracted from the coverage of Arabic, Russian, and French language outlets. However, to fulfil the main purpose of the present work – to demonstrate the roots and applications of the novel concept of 'agenda for action' and its mechanistic principles – it is sufficient to focus on English language news coverage. Therefore, it was decided that the scope of work in the present thesis be limited to English language news outlets. Thus, *The NYT* and *The Guardian* were selected for analysis. This approach has rendered the introduced method somewhat limited, as it is impossible to draw solid and sound conclusions about the situation in Syria at the given time period, based solely on two foreign outlets. By analysing news texts and agendas for action advanced therein retrospectively, it is possible to demonstrate that the policies of the two countries unfold in the way predicted by the outlets. Thus, this approach can be used to track policy by analysing news in real time. The analysis below should, therefore, serve as a demonstration of the interpretation of agendas for action, shedding light onto the problem at hand, but not as a complete and thorough analysis of the Syrian chemical weapon crisis.

The news coverage was retrieved from the archives of *The* NYT and *The Guardian* based on a search for references to Syria (e.g., 'Syria*', 'Damascus') and chemical weapons (e.g., 'chemical weapon', 'sarin', 'WMD', 'conventional weapon', among others). The dataset comprised 572 articles from *The Guardian* and 617 from *The* NYT. Articles were grouped by fifteen-day period, split into sentences, each of which was classified as agenda for action or not, using the trained model described in the third chapter. Thereafter, the extracted agendas were automatically classified, in accordance with the fine-grained classification, into seven labels. The categories were agendas for escalation, agendas for de-escalation, agendas for help and support, agendas for restrictive measures, agendas for not doing, multiclass agendas in one sentence and other (see Section 3.8.2). The detailed statistical analyses of different types of agendas for action over given timeframe in *The* NYT and *The Guardian* are presented in Appendix 4.

4.4 Interpretation of agendas for action extracted from *The NYT* and *The Guardian*

There are multiple ways one can interpret the number and quality of agendas for action expressed in news reports. Below I provide several examples of the manner in which information about agendas for action can be used and interpreted. Note that, in a similar manner, agendas for action can be extracted from other source texts, such as the familiar, social media and political communication texts. While methodology and approach does not change, and no model re-training is required in application to other texts, some examples of interpretation below may be better suited to different text types.

Example 1

Tracking the only volume of agendas for action – without classification – can generate valuable insights about conflict dynamics. Depicting the absolute number and the share of agendas for action expressed across the time, one may note that the number of agendas for action advanced during different time periods – tenser or calmer periods – is different. During times of lull, the number of agendas for action tends to decline, while periods of uneasiness and escalation are characterised by a raise in the expression of agendas for action. This can be seen in Figure 11, which shows the absolute number of sentences classified as agendas for action over a given timeframe, along with the share of such sentences in the analysed texts. As has been shown in the overview of the events (Section 4.1.2), there were two tense periods in 2013 (March and August), when chemical weapon attacks took place. This corresponds to the increase in the share of agendas for action in the graph. The time period from April until early August corresponds to a lull period, with the number of expressed agendas decreasing. The labile property of agendas for action is very useful when analysing conflict events in real time. If the number of agendas for action increases, it might be worth investigating triggers for the change, as it is highly likely that an escalation will follow. Similarly, if the number of agendas for action in news coverage remains stable and categorically low, it may signal de-escalation, or a lull phase of the conflict.

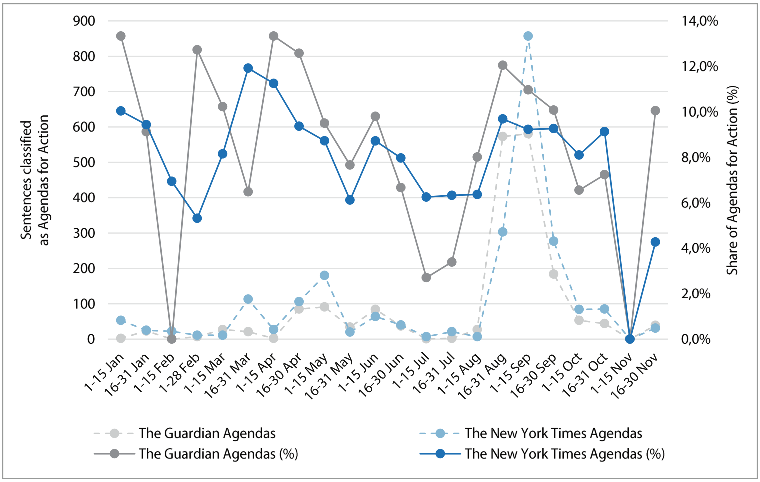


Figure 11: The absolute number and the share of agendas for action advanced in *The NYT* and *The Guardian* from January until November 2013

Example 2

Based on the number of agendas for action advanced in news coverage, one can not only draw conclusions about the situation on the ground, but also gauge the intensity of political debate. In a similar manner to *Example 1*, the increase in the total number of agendas for action in news coverage may reflect heated political discourse, which, in turn, may warrant closer attention to agendas for action advanced by political actors, and the conduct of more detailed analysis thereof. This assumption is proven by retrospective qualitative analysis of political discourse on Syria, its coverage in news outlets, as well as of actual US and UK foreign policy at the time. It is also proven by assessment of the number of agendas for action expressed during the examined time period. Figure 11 shows a steady presence of agendas during the whole coverage period, with two notable crests in both outlets during the second half of March until first half of April, and second half of August through September, respectively. Remarkably, even when the share of agendas for action expressed reached its maximum for the given time period (the first half of April), it was still significantly lower than the average for traditional media, which was recorded as 27.18% (see Table 3).

Syria was the centre of attention for both the UK and the US, the situation there and possible outcomes were discussed by top political actors in both countries multiple times, and from different angles. This interest was reflected by both *The NYT* and *The Guardian*. However, both countries were careful in suggesting possible courses of action, reflected in relatively low numbers of agendas for action (4% to 6 % during lull times). The administration openly canvassed military intervention on the 31st of August reflecting Obama's 'red line' ('Obama warns Syria not to cross 'red line'', 2012), thus surfacing a concrete solution to the problem. This suggestion triggered discussion among supporters and opponents, which, in turn, resulted in an ascending trend in the genesis of agendas for action in *The NYT*, as seen on the graph.

The situation in the UK was very similar to that in the US. Before the vote in the House of Commons on military intervention on the 29th of August, the discussion in political circles became very heated and multiple contesting agendas for action were advanced in order to support or to oppose military actions. Similarly, the shape of the agenda for action curve in *The Guardian* resembles that of *The NYT*. The period of active political discussion corresponds to crests in number of agendas for action. September is characterised by a decrease in number of expressed agendas for action in both outlets. This aligns with the fact that military intervention was no longer seen as an option in both papers; this was due to the negative vote in the House of Commons preventing the UK from proceeding with military attacks against Syria. The UK House of Commons vote became an important factor for the US not to take any military actions either, as their ally – Britain – had deferred from taking military action. Although intervention was no longer considered by either country, the problem of chemical weapons usage remained unsolved. For that reason, alternative solutions were still suggested, including punishment of those responsible for the chemical attacks. After Syria had agreed to eliminate its chemical weapons materials on September 14th, political discussion in both countries became less intense. In Figure 11, one can observe that the second half of September is characterised by decreases in the number of agendas for action advanced in both outlets, which corresponds to a period of calmer political discussion.

In sum, qualitative analysis of actual political discussion and news coverage allows us to draw the conclusion that the number of agendas for action expressed in a news outlet reflects the intensity of political discussion on the topic. This, in turn, enables us to use agendas for action advanced in news coverage as a real-time marker of the state of political discourse. Said marker can serve as a helpful tool, in addition to qualitative analysis of actual state policies. Applied to different textual corpora, such as, parliamentary debates or legislative acts, agendas for action can provide more insights and sound conclusions about state policies.

Example 3
As shown in *Example 2*, analysing agendas for action in news coverage can provide insights into actual foreign policies. Although, using agendas for action extracted from news coverage for this purpose is limited by incomplete or biased presentation of information by news media, it still proves that the tool developed in this thesis can be used as a barometer of polices in real time. Additionally, when applied to a different corpus of texts – such as legislations, press releases, official statements or other news outlets, – the results may provide yet sounder conclusions concerning the content and its outcomes by providing more information thus minimising analysis bias. As a step further, agenda for action can be used to compare foreign political agendas between countries, or to indicate a connection therein (Eder & Kantner, 2000). In accordance with the example above, using news coverage for the purpose of policy comparison can only be achieved to a certain extent. It can, however, serve the purpose of demonstrating usage of the tool. Likewise, qualitative analysis of actual policies, in retrospect, shows that, in the case of the Syrian chemical weapons crisis, *The Guardian* and *The* NYT provided high quality coverage and can in fact be used for policy comparison.

Once more considering Figure 11, during the analysed timeframe the shape of *The Guardian* curve resembles *The* NYT one. Such an observation is especially true over the period from March through November whenever there is a crest or a trough of agendas for action in both outlets, demonstrating similar behaviour. Sometimes not only the trend coincides, but also the number of agendas communicated (see end of

May till beginning of June). This observation may signal impending closure of foreign policies in the two countries regarding the Syrian crisis. It may also signify consensus in potential courses of action. The same is true for the fine-grained classification (see Appendix 4). This observation is especially valuable when tracking policies in real time. In order to draw better founded conclusions and more profound comparisons of foreign policies, agendas for action extracted from other text sources, such as legislations or analytical reports, should be included in further analyses.

Example 4

Tracking the number of agendas for action expressed is enough to offer useful insights into conflict dynamics and policy change. However, it is even more insightful to track the nature of agendas for action being advanced and the change over time of different agendas for action, with respect to actual happenings. As in Examples 1 to 3, analysing news coverage and actual foreign policies of the US and the UK, and comparing them against agendas for action extracted from the news texts, shows that the quantity and quality of agendas for action advanced in news coverage can be used for conflict and policy monitoring in real time, perhaps even gauging future developments. To make more informed assumptions about foreign policies, the agenda for action classifier can be applied to other texts as well, for example speeches or press releases.

Figure 12 plots the distribution of four agenda for action types over the examined time frame: agendas for de-escalation, agendas for escalation, agendas for help and support, and agendas for punishment[18] in *The Guardian*.

18 Here two categories of agendas for action are merged together and are analysed as one class – agendas for punishment and agendas for ignorance. This class is referred to as 'agendas for punishment' in short in this chapter.

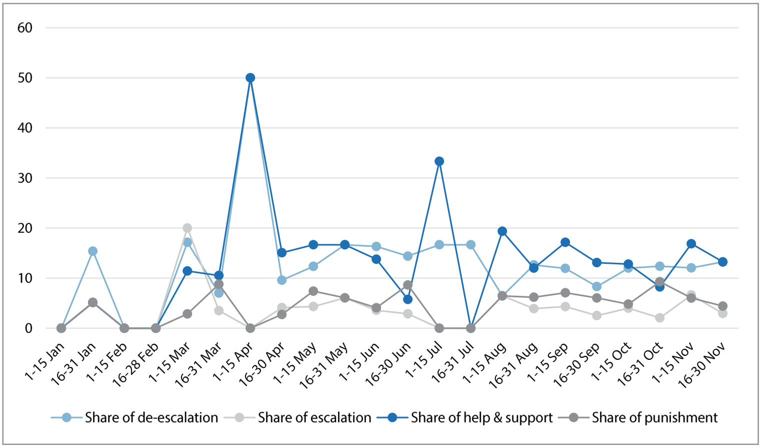


Figure 12: The share of agendas for escalation, de-escalation, help and support and punishment in *The Guardian*

In *The Guardian*, the agenda for escalation was never dominant. Agendas for de-escalation and help and support were the most prominent in the discourse most of the time. Non-violent, non-armed means of handling the conflict were also promoted. The only time when the number of agendas for escalation was slightly higher than agendas for de-escalation was during the first half of March – before attacks in Aleppo. The number of agendas for de-escalation dropped significantly, even equalling the number of agendas for escalation, in the first half of August – prior to the vote in the House of Commons on the 29th of August. When it became clear that the UK would not send forces to Syria together with the détente caused by Syrian chemical disarmament, both agendas for escalation (military intervention) and de-escalation (since the situation in September appeared to stabilize and to be in such a phase) ceased. At that time, agendas for help and support to restore the damaged areas, and to aid victims of the war, became most prominent. Although military intervention was no longer considered, the agenda to punish those accountable for the attack remained relatively high in number. The changes in the number of agendas for escalation and punishment can also be seen in the evolution of dominant frames in the newspaper. The idea of military intervention surfaced already in

March, showing steady presence throughout the investigated period. Even when the UK's participation in the intervention was not considered anymore, the frame to punish the culprit was still quite prominent (Baden & Stalpouskaya, 2015b).

If one looks at the types of agendas for action advanced and plots them over time, in the context of actual decisions made by the UK concerning Syria, it is possible to draw the following conclusion. The quality of agendas for action advanced predicts the direction in which a situation will unfold. In the case of the Guardian, agenda for de-escalation was the dominant one compared to the agenda for escalation, military intervention did not take place during the analysed time period, as the majority of the members of parliament voted against it. If one were to predict the outcome of the vote based on the agenda for action dominating the news coverage, a prediction of no military intervention would have been correct. The same holds for the US media outlet. The agenda for de-escalation was dominating the agenda for escalation during the period examined (see Appendix 4), military intervention did not take place during given timeframe. This property can be used to estimate likelihoods of the possible outcomes of a given situation. It is more likely that the action most ardently called for will actually take place.

Example 5

One of the sources of media agenda – and consequently agendas for action offered by media – is political agenda (van der Pas, 2014; Weaver & Choi, 2017). Hence, policy changes can also be observed through examination of the variety of agendas advanced at different times. This method can also be used to track policy changes in real time. The methodology can also be applied to texts other than news articles.

Figure 13 provides an overview of the structure of agendas for action in *The NYT* over the three most eventful months – July, August and September.

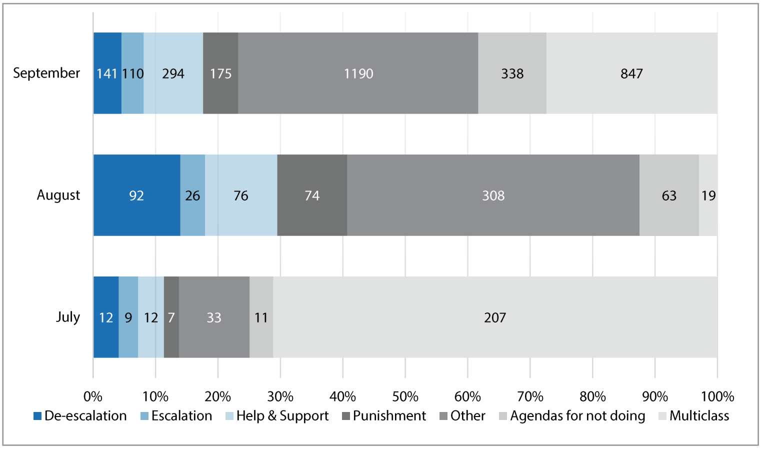


Figure 13: Agenda for Action Landscape in *The NYT* in July, August and September

The dominating agenda for action type in July was 'multiclass' (71%). As mentioned in Section 3.8.2 of the present thesis, sentences that belong to the 'multiclass' agenda for action category consist of multiple clauses, each of them promoting a different action (e.g., one clause calling for escalation, another for de-escalation). The fact that this category was the dominant one in July signals uncertainty and doubts pervading the news discourse. This can be confirmed by qualitative analyses of sentences classified as 'multiclass' agendas for action. For instance, '*But others, particularly many in the State Department, argue that the United States must intervene to prevent a further deterioration of security in the region and to stop a humanitarian crisis that is spiraling out of control, officials said*' (Mazzetti, Schmitt, & Banco, 2013) – such multiclass agendas for action are very typical of *The NYT* in July of the period surveyed. This uncertainty was mainly caused by speculation as to the use of chemical weapons, including the individual or group behind their use. This raised many open questions as to an appropriate means of dealing with the situation in Syria.

The uncertainty which actions to take is in line with the uncertainty frame dominating the outlet at this time as shown by Baden and Stalpouskaya (2015b). This corroborates the claim made in the first chapter of the present thesis that agenda for action serves as operationalisation of treatment recommendation.

The second dominant category of agendas for action in July is 'other', comprising 11% of all agendas for action advanced in this month. As explained in Section 3.8.2, the category 'other' has been introduced due to constraints imposed by computer-assisted analysis, not as an agenda for action category contributing to conflict understanding. Agendas for action with low confidence of classification score are allocated to this category. Given the nature of this category, I would refrain from making conclusions about trends in foreign policy based on these agendas. However, given the context of 'other' agendas for action present in July (high numbers of 'multiclass' agendas and a balanced, though lower, number in the remaining categories) it is reasonable to assume that the large proportion of 'other' agendas for action also signals an uncertain situation, aligned with the governmental consideration of multiple directions of action (see also Section 5.2 for discussion on ambiguity in agenda for action categories). It again fits the dominant uncertainty frame.

The distribution of the remaining classes is balanced at about 4% each. The figures perfectly depict the nature of contemporaneous discussions. Many options were proposed and considered, multiple voices and parties advocate their positions, though the selection of an appropriate course of action was not clear.

The situation became clearer in August after the Ghouta attacks and the decision by the House of Commons against military action. As the most probable ally in case of military intervention deferred, the escalation of the situation became very unlikely, reflected by the low number of agendas for escalation in August – only 4%, the second smallest class after 'multiclass'. However, it was also apparent that the situation could not be left as it was. This is reflected by the increased number of instances in the category of 'agendas for not doing' (most of the examples in this category are sentences like '*The situation cannot be left like this*'). This is also in line with the frame of deterrence gaining dominance at this time. Most agendas for action belong to the categories of 'de-escalation' and 'help and support', which became the most prominent around the same time. Notably, agendas for punishment also comprise a large proportion of advanced agendas – 11%. The context is mostly to punish the party responsible for the use of chemical weapons.

The distribution of agendas for action in September remains similar to that of August. The only difference is that the number of agendas for help and support exceeds agendas for punishment. This matches the notion that Syria had surrendered its chemical stockpiles, thus stating its readiness to cooperate softening the stance of the international community regarding punishment. Using agendas for action to analyse policy changes and development over time can also be used for real-time policy tracking and is especially useful when there is no means of a more qualitative examination; perhaps in such an instance as when the investigator does not speak the language in which information is being delivered. The model must be re-trained to handle texts written in languages other than English (see also Section 5.2) to enhance qualitative analytical capabilities.

4.5 Conclusion

In this chapter, I have demonstrated the interpretation of agendas for action and used the automated tool for their extraction and classification. This endeavour had multiple purposes, including tracking the situation on the ground in conflict areas, predicting its resolution, monitoring foreign policies in real time and comparing policies of different countries, as well as tracking policy changes over time. It may also be possible to predict the outcome of evens, based on agendas for action advanced in texts. I have chosen the coverage of chemical weapon crisis in Syria in 2013 in *The NYT* and *The Guardian* to showcase these capabilities. The observed period features high levels of uncertainty and a large number of contradicting agendas for action. For this reason, it is a very suitable showcase of the use and interpretation of agendas for action. Given that only two newspapers, and no sources that would have been more suitable for policy analysis (including legislations, official statements, and talks), were used for this study, the results and findings presented in this chapter remain somewhat limited. The main purpose of these findings is to demonstrate possible usage and capacities of the tool developed. The ways agendas for action have been used in present thesis are suggestions and examples of possible analytical approaches. Depending on the goal, there may be other ways to comprehend them. Even with some limitations, such as relatively low classification scores and the inability to analyse non-English language media, the tool used to analyse agendas for action proves powerful, capable of fulfilling all these tasks.

5 Contributions, limitations and future work

Throughout this dissertation, I have introduced the novel concept of 'agenda for action', which advances multiple branches of communication science and linguistics. I have also developed an algorithm which enables automatic extraction and classification of agendas for action, also demonstrating its use in monitoring conflict dynamics on the ground. I have also shown that the algorithm can be used to compare news coverage and foreign policies in real time. In this final chapter I summarise the main findings and contributions of this dissertation as well as outline the limitation of the work. I end by offering some directions for future research.

5.1 Contributions

The contributions of the current dissertation are manifold. Being an interdisciplinary work, it advances various scholarly approaches. The present thesis bridges and fills in the missing pieces of three related theories in communication science – framing, agenda setting and collective memories. These theories discuss public agendas from different angles: Agenda setting investigates issues, framing focuses on attitudes towards them, and the collective memories approach studies complex phenomena that exist in the consciousness of peoples, with respect to the past (the memories) and the future (intentions), the latter being embedded in agendas for action. Not only does the concept of 'agenda for action' developed in the present work touch upon all of these theories, it can also further advance them and fill in existing gaps therein. It extends the well-established sequence of agenda properties: Agenda setting defines what to think *about*, framing decides *how* to think about it and, finally, agenda for action prescribes what to *do* concerning these issues. Not only does the framework of 'agenda for action' bridge three theoretical frameworks together, but it also fills the term 'agenda' with more concrete meaning and substantial understanding. In fact, classical agenda setting theory has focussed on a topic, an issue, what the content of talk and thought, lacking agenda itself. The term 'agenda for action' (which etymologically might sound redundant) adds the missing 'action' component to agenda setting. It also explains the process of agendas being set and brought to life as an action upon an issue or thought. This account has been only crudely explained in agenda setting and briefly mentioned in framing.

Introducing the concept of 'agenda for action', is evident that public agenda is set via the expression of an illocutionary force and communicated to a hearer (see Section 1.2.3). This is realised with the help of linguistic constructs called 'speech acts'. They can be communicated and perceived directly, when, for instance, the audience is listening to a live broadcast or reading a transcript. However, most frequently, the process of agenda setting is mediated by journalists who report and frame agendas expressed by the élite (Kampf, 2013). This creates an additional level of pragmatic meaning for an utterance. Speech acts are traditionally context-dependent and perform an illocutionary act only when certain conventions are met (Roberts, 2017; Sbisà, 2002; van Dijk, 1977). Mediated speech acts follow different rules and conventions and must be interpreted in a different context. The present study integrates speech act theory into the framework of communication studies, demonstrating and explaining that agendas are being set through the media of speech acts. The current thesis has also made advances in the theory of speech acts by emphasising the direction of fit as their key criterion and basis for classification (Roberts, 2017). It unites direct, indirect, explicit and implicit speech acts with the direction of fit from word to world, within the concept of 'agenda for action'. The present thesis also contributes to the study of collective speech acts (Meijers, 2007), as agenda for action either addresses multiple hearers or are expressed by a group (i.e. a state, an organisation, etc.), thus representing collective speech acts.

Alongside theoretical contributions, this dissertation offers several methodological advances. First of all, it brings the methodological framework of communication science forward by operationalising the concept of 'treatment recommendation'. Among the four prime frame components suggested by Entman (1993) treatment recommendation embedded in agendas for action has been the least operationalised. This thesis provides a sound theoretical basis to identify and interpret agendas for action thus bringing a new dimension, as well as the tool for its measurement to the toolkit of classical content analysis. This work also augments the methods of communication science with NLP (natural language processing) approaches to text analysis (e.g., part-of-speech tagging, grammar parsing) and statistical learning through the application of machine learning algorithms in the communication science domain. Finally, the present work also introduces a new type of information to be extracted and analysed from texts. Computational linguistics has been mainly focusing on extracting rather formal low-level phenomena such as parts of speech and syntactic structures, or fulfilling tasks such as machine translation or topical classification of texts. An agenda for action is a complex, high-level semantic concept which takes multiple dimensions into account; grammar, meaning, and context. Combining existing tools (e.g., grammar parsers and lemmatisers, see Section 2.1.3) one can build an algorithm capable of extracting and analysing complex phenomena which has been extensively showcased in the current work.

To demonstrate how these advances can be used practically, an algorithm based on machine learning has been developed. It is capable of considering both low- and high-level information and distilling complex semantic entities – agendas for action – from news coverage. As explained in Section 3.2, the algorithm uses bag-of-words and n-gram statistics, as well as lexical, morphological and syntactic information to classify propositions as agendas for action, as well as assigning them to a given class. The code is open source and can be used without any limitations.[19]

Machine learning has proven more powerful as an approach than hand-crafted rules when dealing with big data, including heterogeneous texts and complex phenomena such as agendas for action. However, it is only possible to use machine learning when extensive data are available. In the age of the Internet, it is fairly easy to gather data, especially samples of mediated communication. However, in order to perform supervised machine learning (which is the case with agendas for action), the data need to be appropriately annotated. Developing

19 https://github.com/KatStal/a4a_extractor corpora that can be used for machine learning is extremely tedious. There exists no corpus that can be used for the purpose of the present thesis, and so a corpus of agendas for action had to be compiled. This is another of the main contributions of this work. It has led to the creation of the first corpus of agendas for action, which may be provided for research purposes upon request.

The present work also shows in several ways that the ability to extract and classify agendas for action from texts about war and conflict can help to track the development of conflict dynamics and national policies. I have demonstrated the analysis of agendas expressed in *The* NYT and *The Guardian* over the course of the Syrian chemical weapons crises in 2013 reflects the political agenda prevailing in media discourse in the USA and the UK. Furthermore, it can be seen that the dominant agendas for action are rooted in the collective conscious, leading to certain courses of collective actions. For example, I have demonstrated that the agenda for escalation and military solution to the crisis was never the most prominent in either news outlet. Indeed, it is known the crisis in 2013 was solved by diplomatic means and military intervention did not take place, though tracking agenda for action in real-time may have predicted this outcome. This is a first indication that tracking the amount and the quality of agendas for action in news coverage may predict collective action and mobilisation of force, foresing the outcome of a multifactorial situation.

5.2 Limitations and future work

The abovementioned theoretical and methodological advances open a broad field for future research and application of the concept of 'agenda for action' developed in this work. Here, I have only shown one possible policy domain where agendas for action could provide useful insights. However, as a universal way to advance policy, the concept of 'agenda for action' can help to track political transition in any domain, including financial reform, migration law, even climate change. One of the main advantages of the method suggested in this thesis is the possibility to extend it use to other discourse styles and genres. The ability to track policy change in real time is especially valuable. Extracting agendas for action from different media and in different languages immediately following publication and displaying the changes as dashboards (similar to Figure 12) could serve as something of a social barometer. It can help to observe the dynamics of policy change and advanced agendas, identifying potentially dangerous trends in a timely manner, leading to their proactive management. Comparing and tracking agendas for action advanced at the same time, but in different media arenas (local and international), could also provide useful insight and deeper understanding of real world events contemporaneously. Moreover, as calling for action is one of the prime functions of language and communication (Jakobson, 1960), agendas for action appear in texts of different genres and topics. For this reason, the algorithm introduced in the current work can be extended to other texts, not only news coverage. For instance, one of the possible applications of the agendas for action extractor could be the analysis of emails and generation of a to-do lists based on their content. Another application of agendas for action would be the analysis of medical texts and classification of prescriptions: medicine to take, treatments to undergo, and nutritional recommendations.

In order to extend the approach and the algorithm developed in this thesis, several limitations need first to be addressed. To extract and analyse agendas for action in domains other than war and violent conflict, a respective corpus must be crafted and annotated accordingly. While extracting agendas for action from texts (the first step is discerning agendas versus not agendas) is universal and can be applied to texts regardless of source and genre, the second step, fine-grained classification, is domain specific and cannot be used to extract agendas from texts of different domains. The property of language to call for something is embedded mainly in grammar or domain unspecific words (such as the verbs 'encourage', 'ask for', 'command'). On the contrary, fine-grained classification, such as in the present dissertation, is domain specific. This means that most classes extracted and analysed by the developed tool can be only found in war and conflict related texts. To analyse agendas for action in texts of different topics, a new agenda for action taxonomy would need to be developed. When analysing texts about climate change and environmental policies, the following classification of agendas for action would make sense: agendas for stricter emission control policies, agendas for higher taxes for hazardous industries, agendas to lower the emission quotas, and agendas to subsidise enterprises that use renewable energy. After a respective corpus has been compiled and annotated, the classifier must be re-trained as well, since it was trained on the corpus annotated for agendas of action in war and violent conflict. The second step of classification is very specific and needs to be individually tailored to a given classification task. Since data preparation is in fact the most time- and labour-intensive part of the process, the suggested approach of supervised learning remains limited. Application of unsupervised learning techniques, such as clustering, might be worth exploration in order to optimize the fine-grained classification step (Hastie et al., 2001).

The same holds for analysing media in languages other than English. A language-specific corpus needs to be compiled. Without knowledge of the language, it is extremely hard to follow the narrative in national and local news outlets and comprehend content addressing local people. However, these sources of information can be extremely relevant and interesting, as they reflect the most immediate moods, policy trends, and changes at the coalface. For this case, agendas for action can also serve as a real time measure of political moods and intentions, notifying of a change in the course of action, thereby enabling third parties to act in accordance with new developments. Being able to quickly grasp dominant agendas in a discourse may further enable analysis and comparison of the same events through the prism of local and international media, facilitating the observation of policy on a national and international level.

Many improvements can be made to the algorithm itself. The classifier confidence level is low, especially when concerned with fine-grained classification. It often confuses agendas of different classes (e.g., punishment and ignorance). Predictions made by a classifier with the accuracy score of 50% are as good as chance. In case of accuracy lower than 50%, chance would yield more confident classification decisions. To ameliorate the problem of low accuracy, more training data is needed, especially for underrepresented classes. Enlarging the corpus should be one of the top priorities for improving future research with the current approach. Linguistic features can be further enhanced by extending lists of hint words and refining disambiguation criteria. Then, new ways to represent data might be explored. For example, it might be worthwhile to pursue word embeddings that are now standard in deep learning (Mikolov et al., 2013). Moreover, given more training data and presenting them as word vectors would enable the use of advanced algorithms such as neural networks, which, in turn, would eliminate the need to further hand-craft linguistic features. These processes would render extraction and classification of agendas for action a purely data-driven endeavour.

The final step of the current implementation is classifying agendas for action. Their interpretation and subsequent use remain up to a researcher. Being a leap forward in content analysis and providing many insights, agendas for action are even more valuable if examined in context. For instance, when linked to their utterer or to the addressee, it is easier to make sense of agendas and of upcoming changes in policy (van Atteveldt, Sheafer, Shenhav, & Fogel-Dror, 2017). When the source or the author of an agenda for action is known, this limitation appears to be a minor problem. This is true for texts produced by specific actors, such as press releases, policy drafts, or speeches by political parties or NGOs, when the source of an agenda is previously established. Similarly, automated agenda for action extraction can be applied to texts stemming from social media, where the utterer can be tracked. It is possible that actors such as politicians or NGOs refer to or share posts of others, thus agendas that are contained therein will have different sources. For such cases, the ability to link an agenda to its origin would be helpful to the interpretation of meaning. However, social media posts are mainly shared in order to support and strengthen one's own purview, hence the owner of the social media account is most likely to re-post information that is pragmatically close to their own standpoint (Larsson & Moe, 2011; Stieglitz & Dang-Xuan, 2013). On the contrary, when analysing news coverage, it is not enough to know which outlet has published an agenda for action, it is more important to know who owns the agenda, as newspapers tend to advance an established agenda, rather than extol a new one. It is also important to know whether the quoted agenda was endorsed or disapproved. Thus, the same article may contain contradicting agendas, as long as they are promoted by different parties, for example, *'Russia wants investigators to examine only accusations by each side that the other used chemical weapons in Aleppo on March 19. Britain, France and the United States urge the inclusion of other opposition allegations that Mr. Assad's government used chemical weapons in Damascus and Homs'* (Droubi & Gladstone, 2013). Attributing an agenda for action to a speaker in an automated fashion is feasible. One needs to perform grammatical parsing of a sentence expressing an agenda (or of a sentence) and attribute it to a source. However, this step has not been performed in this research.

Another thing that can be re-worked in current approach is the agenda for action taxonomy. The one developed in this thesis involves vague and broad categories such as 'other' and 'multiclass', which are hard to understand and interpret. The category 'other' is a catch-all for hard-to-classify cases. It is also often the category where misclassified agendas land. 'Other' agendas for action are hard to interpret as advanced agendas may be 'positive', as well as 'negative', thus the fact that there are 'other' agendas present in a text does not facilitate clear understanding of the situation. When interpreting the results of automated classification, the category 'other' seems to be rather useless, relatively unable to predict the development of a conflict. As has been demonstrated in previous chapters, the increased number of 'other' agendas for action may signal rising uncertainty, but, to be able to draw conclusions based on the insights from the analysis, the 'other' agendas need to be de-coded and aligned with meaningful classes in the taxonomy.

Analogously, the category 'multiclass' poses the same challenges. The category includes those sentences that call for multiple, sometimes contradictory, actions, e.g., *'The letter says governments supporting the opposing sides in the civil war should use their influence to stop the attacks and the un and international donors must do more to increase support to Syrian medical networks'* (Siddique, 2013) – in this example there are two agendas for action. The former is an agenda for not doing, while the latter is an agenda for help and support. However, in accordance with the current taxonomy, the sentence falls into the 'multiclass' category. Although, it is valuable to know that agendas expressed in a text are multiplex, potentially signalling uncertainty and turbulence (as in the Syrian case), to thoroughly understand the political landscape and make well-informed assumptions about appropriate activity, each agenda for action within such complex sentences must be considered individually. Technically, complex sentences that express multiple agendas must be parsed into clauses, and each clause ought to be analysed and classified separately. To implement this requirement, one would need to perform and additional step in the pre-processing stage, splitting all sentences from the input text set into clauses. This step should be performed in future studies, and would further enhance the capabilities of the algorithm.

5.3 Conclusion

This doctoral dissertation has been written to serve several purposes. My first intention was to produce an interdisciplinary work covering communication science, linguistics and computational linguistics. On the one hand, it aims at filling in some existing gaps in the theoretical and methodological frameworks of the three disciplines. On the other hand, it builds a bridge between these fields. Agendas for action as requests for activity are a natural and universal property of interpersonal communication. Automatic extraction of agendas for action is a powerful and useful tool for understanding violent conflict and its coverage in news media. It enables the automatic identification of the content of a call to action, thus allowing for more informed assumptions about potential outcomes and their appropriate reactions. Even though the scholarship has touched upon agendas for action in many ways, it remains under-researched and offers much room for investigation. I hope that this work has sketched a course for future directions in the field, for those interested in this ubiquitous, yet complex phenomenon. Moreover, I believe that an interdisciplinary approach to agendas for action will produce a thorough understanding, contributing to many fields of science. Potential applications and ideas for future research outlined above are some suggestions, I would be delighted to see how the ideas and findings provided in this work are challenged and adapted by future scholars.

It is my dream that the idea of agendas for action as a social barometer and monitor of policy changes in real-time may be adapted to many applications. I further hope that the tool developed in this study finds a home in the toolbox of scholars, and indeed anyone, interested in this topic. I am very curious to see the classifier used in different domains and for different purposes than news analysis.

This thesis differs from more standard PhD projects, where clearly stated research questions get answered in the course of the work. My thesis neither asks nor answers any research questions. In fact, it might leave the disciplines with even more open questions than when I commenced my work. I do believe, however, that the theoretical and methodological advances presented here will inspire fellow researchers, leading to the crystallisation of answers to many questions from many scholarly traditions.

References

A Timeline of the Syrian conflict as it enters its eighth year (2018, March 15). *Bloomberg*. Retrieved from https://www.bloomberg.com/news/articles/2018-03-15/timeline-of-the-syrian-conflict-as-it-enters-8th-year

Abu-Mostafa, Y. S., Magdon-Ismail, M., & Lin, H.-T. (2012). *Learning from data: A short course*. S.l.: AMLbook.com.

Aggarwal, C. C., & Zhai, C. (Eds.). (2012). *Mining text data*. Springer US. http://dx.doi.org/10.1007/978-1-4614-3223-4

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), *Springer series in social psychology. Action control: From cognition to behavior* (1st ed., pp. 11–39). Springer. https://doi.org/10.1007/978-3-642-69746-3_2

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior* (Pbk. ed.). Prentice-Hall.

Ali, M., Khalid, S., Rana, M. I., & Azhar, F. (2018). A probabilistic framework for short text classification. In *2018 IEEE 8th annual computing and communication workshop and conference (CCWC)* (pp. 742–747). https://doi.org/10.1109/CCWC.2018.8301712

Al-Rawi, A. K. (2015). Sectarianism and the Arab Spring: Framing the popular protests in Bahrain. *Global Media and Communication*, *11*(1), 25–42. https://doi.org/10.1177/1742766515573550

Anheier, H. K., Neidhardt, F., & Vortkamp, W. (1998). Movement cycles and the Nazi party: Activities of the Munich *NSDAP*, 1925–1930. *American Behavioral Scientist*, *41*(9), 1262–1281. https://doi.org/10.1177/0002764298041009006

Anscombe, G. E. M. (1957). *Intention*. Basil Blackwell.

Arrese, Á., & Vara-Miguel, A. (2015). A comparative study of metaphors in press reporting of the Euro crisis. *Discourse & Society*, *27*(2), 133–155. https://doi.org/10.1177/0957926515611552

Artstein, R., & Poesio, M. (2008). Inter-coder agreement for computational linguistics. *Computational Linguistics*, *34*(4), 555–596. https://doi.org/10.1162/coli.07-034-R2

Austin, J. L. (1962). How to do things with words: [the William James lectures delivered at Harvard University in 1955] (2. ed., [repr.]). Harvard Univercity Press.

Bach, K. (1994). Conversational implicature. *Mind & Language*, *9*(2), 124–162. https://doi.org/10.1111/j.1468-0017.1994.tb00220.x

Bach, K. (2014). Speech acts. In *Routledge encyclopedia of philosophy.* Retrieved from http://online.sfsu.edu/kbach/spchacts.html

Bach, K., & Harnish, R. M. (1979). Linguistic communication and speech acts. *The Philosophical Review*, *91*(1), 134. https://doi.org/10.2307/2184680

Baden, C. (2010). Communication, contextualization, & cognition: Patterns & processes of frames' influence on people's interpretations of the EU constitution (Doctoral dissertation). Eburon. Retrieved from https://pure.uva.nl/ws/files/1036231/80640_thesis.pdf

Baden, C., & Stalpouskaya, K. (2015a). Common methodological framework: content analysis. A mixed-methods strategy for comparatively, diachronically analyzing conflict discourse. INFOCORE Working Paper 10/2015. INFOCORE. Retrieved from http://www.infocore.eu/wp-content/uploads/2016/02/Methodological-Paper-MWG-CA_final.pdf

Baden, C., & Stalpouskaya, K. (2015b, May). Maintaining frame coherence between uncertain information and changing agendas: The evolving framing of the Syrian chemical attacks in the US, British, and Russian news. Paper presented at the ICA Annual Conference.

Baden, C., & Tenenboim-Weinblatt, K. (2016). *Agenda for action: INFOCORE definitions*. Retrieved from http://www.infocore.eu/wp-content/uploads/2016/02/def_agendas_for_action.pdf

Balakrishnan, V., & Ethel, L.-Y. (2014). Stemming and lemmatization: A comparison of retrieval performances. *Lecture Notes on Software Engineering*, *2*(3), 262–267. https://doi.org/10.7763/LNSE.2014.V2.134

Balanche, F. (2018). *Sectarianism in Syria's civil war*. Retrieved from https://www.washingtoninstitute.org/policy-analysis/view/sectarianism-in-syrias-civil-war

Bavelas, B. J., Black, A., Chovil, N., & Mullett, J. (1990). Equivocal communications. *Sage Series in Interpersonal Communication*, *11*.

Bayes, T., & Price, R. (1763). An essay towards solving a problem in the doctrine of chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. *Philosophical Transactions of the Royal Society of London*, *53*(0), 370–418. https://doi.org/10.1098/rstl.1763.0053

Becker, L. B. (1977). The impact of issue saliences. In D. L. Shaw & M. E. McCombs (Eds.), *The emergence of American political issues: The agenda-setting function of the press* (pp. 121–131). West Publishing Co.

Benford, R. D., & Snow, D. A. (2000). Framing processes and social movements: An overview and assessment. *Annual Review of Sociology*, *26*(1), 611–639. https://doi.org/10.1146/annurev.soc.26.1.611

Berelson, B. (1996). KAP studies on fertility. In B. Berelson, R. K. Anderson, O. Harkavy, J. Maier, W. P. Mauldin, & S. J. Segal (Eds.), *Family planning and population programs: A review of world developments* (pp. 665–668). University of Chicago Press.

Bertolet, R. (1994). Are there indirect speech acts? In S. L. Tsohatzidis (Ed.), *Foundations of speech act theory: Philosophical and linguistic perspectives* (pp. 335–349). Routledge.

Bimber, B. (2017). Three prompts for collective action in the context of digital media. *Political Communication*, *34*(1), 6–20. https://doi.org/10.1080/10584609.2016.1223772

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python. Safari Books Online*. O'Reilly Media Inc. Retrieved from https://www.oreilly.com/library/view/natural-language-processing/9780596803346/

Bowe, B. J. (2014). *Framing moral evaluations: Moral foundations in U.S. newspaper coverage of mosque controversies* (Doctoral dissertation). ProQuest Dissertations Publishing. Retrieved from https://cedar.wwu.edu/cgi/viewcontent.cgi?article=1002&context=cedarbooks

Bowe, B. J., Oshita, T., Terracina-Hartman, C., & Chao, W.-C. (2014). Framing of climate change in newspaper coverage of the East Anglia e-mail scandal. *Public Understanding of Science (Bristol, England)*, *23*(2), 157–169. https://doi.org/10.1177/0963662512449949

Brown, P., & Levinson, S. C. (2011[1987]). Politeness: Some universals in language usage (21. pr). Studies in interactional sociolinguistics: Vol. 4. Cambridge University Press.

Bryant, J., & Miron, D. (2004). Theory and research in mass communication. *Journal of Communication*, *54*(4), 662–704. https://doi.org/10.1111/j.1460-2466.2004.tb02650.x

Burscher, B., Odijk, D., Vliegenthart, R., Rijke, M. de, & Vreese, C. H. de. (2014). Teaching the computer to code frames in news: Comparing two supervised machine learning approaches to frame analysis. *Communication Methods and Measures*, *8*(3), 190–206. https://doi.org/10.1080/19312458.2014.937527

Cabrera, J. M., Escalante, H. J., & Montes-y-Gomez, M. (2013). Distributional term representations for short-text categorization. *Computational Linguistics and Intelligent Text Processing*, 335–346. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.386.5397&rep=rep1&type=pdf

Card, D., Boydstun, A. E., Gross, J. H., Resnik, P., & Smith, N. A. (2015). The media frames corpus: Annotations of frames across issues. In Min-Yen Kan (Ed.), *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)* (pp. 438–444). Retrieved from https://www.aclweb.org/anthology/P15-2072.pdf

Carvalho, V. R., & Cohen, W. W. (2005). On the collective classification of email 'speech acts'. In *Proceedings of the 28th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 345–352). ACM. Retrieved from http://www.cs.cmu.edu/~vitor/papers/sigir05.pdf

Carvalho, V. R., & Cohen, W. W. (2006). Improving 'email speech acts' analysis via n-gram selection. In *Proceedings of the HLT-NAACL 2006 workshop on analyzing conversations in text and speech, ACTS '09* (pp. 35–41). Association for Computational Linguistics. Retrieved from https://www.aclweb.org/anthology/W06-3406.pdf

Cavnar, W. B., & John M. Trenkle (1994). N-Gram-based text categorization. In *Proceedings of SDAIR* (pp. 161–175). Retrieved from https://www.let.rug.nl/~vannoord/TextCat/textcat.pdf

Chong, D., & Druckman, J. N. (2007). Framing theory. *Annual Review of Political Science*, *10*(1), 103–126. https://doi.org/10.1146/annurev.polisci.10.072805.103054

Cohen, T. (1973). Illocutions and perlocutions. *Foundations of Language*, *9*(4), 492–503.

Cohen, W. W., Carvalho, V. R., & Mitchell, T. M. (2004). Learning to classify email into 'speech acts'. *In Proceedings of the 2004 conference on empirical methods in natural language processing* (pp. 309–316). Retrieved from https://www.aclweb.org/anthology/W04-3240.pdf

Colas, F., & Brazdil, P. (2006). Comparison of SVM and some older classification algorithms in text classification tasks. In M. Bramer (Ed.), *IFIP International federation for information processing: Vol. 217. Artificial intelligence in theory and practice: IFIP 19th world computer congress, TC 12: IFIP AI 2006 Stream* (Vol. 217, pp. 169–178). International Federation for Information Processing. https://doi.org/10.1007/978-0-387-34747-9_18

Conway, D., & White, J. M. (2012). *Machine learning for hackers: [case studies and algorithms to get you started]* (1. ed.). O'Reilly. Retrieved from https://doc.lagout.org/science/Artificial%20Intelligence/Machine%20learning/Machine%20Learning%20for%20Hackers_%20Case%20Studies%20and%20Algorithms%20to%20Get%20You%20Started%20%5BConway%20%26%20White%202012-02-25%5D.pdf

Denham, B. E. (2010). Toward conceptual consistency in studies of agenda-building processes: A scholarly review. *Review of Communication*, *10*(4), 306–323. https://doi.org/10.1080/15358593.2010.502593

Dobbin, K. K., & Simon, R. M. (2011). Optimally splitting cases for training and testing high dimensional classifiers. *BMC Medical Genomics*, *4*, 31. https://doi.org/10.1186/1755-8794-4-31

Donk, A., Metag, J., Kohring, M., & Marcinkowski, F. (2012). Framing emerging technologies: Risk perceptions of nanotechnology in the German press. *Science Communication*, *34*(1), 5–29. https://doi.org/10.1177/1075547011417892

Droubi, H., & Gladstone, R. (2013, March 26). Syrian opposition joins meeting of Arab League. *The New York Times*. Retrieved from https://www.nytimes.com/2013/03/27/world/middleeast/syrian-opposition-group-takes-seat-at-arab-league.html

Dumais, S., Platt, J., Heckerman, D., & Sahami, M. (1998). Inductive learning algorithms and representations for text categorization. In N. Pissinou (Ed.), *Proceedings of the seventh international conference on Information and knowledge management* (pp. 148–155). ACM. https://doi.org/10.1145/288627.288651

Eder, K., & Kantner, C. (2000). Transnationale Resonanzstrukturen in Europa: Eine Kritik der Rede vom Öffentlichkeitsdefizit. In M. Bach (Ed.), *Die Europäisierung nationaler Gesellschaften* (pp. 306–331). Westdeutscher Verlag.

Ellis, J. (1996). Prospective memory or the realization of delayed intentions: A conceptual framework for research. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 1–22). Lawrence Erlbaum Associates Publishers.

Ellis, J., & Cohen, G. (2008). Memory for intentions, actions and plans. In G. Cohen & M. A. Conway (Eds.), *Memory in the real world* (3rd ed., pp. 141–172). Psychology Press.

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, *43*(4), 51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x

Entman, R. M. (2003). Cascading activation: Contesting the White House's frame after 9/11. *Political Communication*, *20*(4), 415–432. https://doi.org/10.1080/10584600390244176

Fan, X., & Hu, H. (2010). A new model for Chinese short-text classification considering feature extension. In *2010 international conference on artificial intelligence and computational intelligence (AICI)* (pp. 7–11). https://doi.org/10.1109/AICI.2010.125

Fillmore, C. J. (1976). Frame semantics and the nature of language. *Annals of the New York Academy of Sciences*, *280*(1), 20–32. https://doi.org/10.1111/j.1749-6632.1976.tb25467.x

Fishbein, M. (1967). Attitude and prediction of behavior. In M. Fishbein (Ed.), *Readings in attitude theory and measurement* (pp. 477–492). Wiley.

Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention and behavior: An introduction to theory and research. Addison-Wesley series in social psychology. Addison-Wesley.

Fröhlich, R. (Ed.). (2018). Public communication of war and armed conflict: Dynamics of conflict news. [S.l.]. Routledge.

Fürnkranz, J. (1998). A study using n-gram features for text categorization: Technical report.

Galtung, J. (2013). High road, low road: Charting the course for peace journalism. In J. Galtung & D. Fischer (Eds.), *Springer briefs on pioneers in science and practice: Vol. 5. Johan Galtung: Pioneer of peace research* (Vol. 5, pp. 95–102). Springer. https://doi.org/10.1007/978-3-642-32481-9_8

Gamson, W. A. (2005). Robert M. Entman. Projections of power: Framing news, public opinion, and U.S. foreign policy. *Public Opinion Quarterly*, *69*(2), 324–326. https://doi.org/10.1093/poq/nfi017

Gamson, W. A. (1992). The social psychology of collective action. In A. D. Morris & C. M. Mueller (Eds.), *Frontiers in social movement theory* (pp. 53–76). Yale University Press.

Gamson, W. A. (1995). Constructing social protest. In H. Johnston & B. Klandermans (Eds.), *Social movements and culture* (pp. 85–106). University of Minnesota Press.

Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural Networks and the Bias/Variance Dilemma. *Neural Computation*, *4*(1), 1–58. https://doi.org/10.1162/neco.1992.4.1.1

Gerhards, J., & Rucht, D. (1992). Mesomobilization: Organizing and framing in two protest campaigns in West Germany. *The American Journal of Sociology*, *98*(3), 555–596.

Ghanem, S. (1997). Filling in the tapestry: The second level of agenda-setting. In M. E. McCombs, D. L. Shaw, & D. H. Weaver (Eds.), *Communication and democracy: Exploring the intellectual frontiers in agenda-setting theory* (pp. 3–14). Lawrence Erlbaum Associates.

Ghorpade, S. (1986). Agenda setting: A test of advertising's neglected function. *Journal of Advertising Research*, *26*(4), 23–27.

Giugni, M. (2006). *Problem framing and collective action in the field of unemployment*. Paper presented at the International conference on contentious politics and social movements in the 21th century.

Goffman, E. (1974). Frame analysis: an essay on the organization of experience. Harper & Row.

Golan, G., & Wanta, W. (2001). Second-level agenda setting in the New Hampshire primary: A comparison of coverage in three newspapers and public perceptions of candidates. *Journalism & Mass Communication Quarterly*, *78*(2), 247–259. https://doi.org/10.1177/107769900107800203

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. Retrieved from https://www.deeplearningbook.org

Goretti, L. N. (2007). Peace journalism applied: An assessment of media coverage of the conflict in Northern Uganda. *Conflict & Communication Online*, *6*(2). Retrieved from http://www.cco.regener-online.de/2007_2/pdf/nassanga.pdf

Green, M. (1999). Illocutions, implicata, and what a conversation requires. *Pragmatics & Cognition*, *7*, 65–92.

Greene, D., & Cross, J. P. (2016). Exploring the political agenda of the European parliament using a dynamic topic modeling approach. Retrieved from http://arxiv.org/pdf/1607.03055

Gwet, K. L. (2014). Handbook of inter-rater reliability (4th edition): The definitive guide to measuring the extent of agreement among raters. Advanced Analytics, LLC.

Hale, J. L., Householder, B. J., & Greene, K. L. (2002). The theory of reasoned action. In J P. Dillard & M. Pfau (Eds.), *The persuasion handbook: developments in theory and practice* (pp. 259–286). SAGE Publications. http://dx.doi.org/10.4135/9781412976046

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software. *ACM SIGKDD Explorations Newsletter*, *11*(1), 10. https://doi.org/10.1145/1656274.1656278

Hanitzsch, T. (2004). Journalists as peacekeeping force? Peace journalism and mass communication theory. *Journalism Studies*, *5*(4), 483–495. https://doi.org/10.1080/1461670042331296419

Hardy, A., & Sevenans, J. (2015, May). *The political agenda-setting effect of media storms: A story-level analysis.* Paper presented at the ICA Annual Conference.

Harnish, R. M. (1994). Mood, meaning and speech acts. In S. L. Tsohatzidis (Ed.), *Foundations of speech act theory: philosophical and linguistic perspectives* (pp. 407–459). Routledge.

Harris, S. (2001). Being politically impolite: extending politeness theory to adversarial political discourse. *Discourse & Society*, *12*(14), 451–472. https://doi.org/10.1177%2F0957926501012004003

Hassan, S., Rafi, M., & Shaikh, M. S. (2011). *Comparing SVM and naïve Bayes classifiers for text categorization with Wikitology as knowledge enrichment*. Paper presented at the IEEE 14th International Multitopic Conference.

Hastie, T., Friedman, J., & Tibshirani, R. (2001). *The elements of statistical learning: Data Mining, inference, and prediction*. *Springer series in statistics*. Springer New York. Retrieved from http://dx.doi.org/10.1007/978-0-387-21606-5

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*(1), 77–89. https://doi.org/10.1080/19312450709336664

Helfer, L., & Wonneberger, A. (2015). Bowling alone, watching together? Structural explanations of social TV viewing from 1990 to 2010. Paper presented at the ICA Annual Conference.

Hofmann, T. (1999). Probabilistic latent semantic analysis. In K. B. Laskey (Ed.), *Uncertainty in artificial intelligence: Proceedings of the fifteenth conference* (pp. 289–296). Morgan Kaufmann.

Holdcroft, D. (1994). Indirect speech acts and propositional content. In S. L. Tsohatzidis (Ed.), *Foundations of speech act theory: Philosophical and linguistic perspectives* (pp. 350–364). Routledge.

Holliday, J. (2011). *Struggle for Syria in 2011: An operational and regional analysis*: Middle East security report 2. Retrieved from http://www.understandingwar.org/sites/default/files/Struggle_For_Syria.pdf

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. In *Proceedings of the National Academy of Sciences of the USA*, (pp. 2554–2558). MIT Press. https://doi.org/10.1073/pnas.79.8.2554

Hu, X., Sun, N., Zhang, C., & Chua, T.-S. (2009). Exploiting internal and external semantics for the clustering of short texts using world knowledge. In *Proceedings of international conference on information and knowledge management*. https://doi.org/10.1145/1645953.1646071

Hu, Y. J., Jiang, J. X., & Chang, H. Y. (2013). Chinese short-text classification based on topic model with high-frequency feature expansion. *Journal of Multimedia*, *8*(4), 425–431. https://doi.org/10.4304/jmm.8.4.425-431

Humberstone, L. (1992). Direction of fit. *Mind*, *101*(401), 59–83.

Iyengar, S. (1990). The accessibility bias in politics: Television news and public opinion. *International Journal of Public Opinion Research*, *2*(1), 1–15. https://doi.org/10.1093/ijpor/2.1.1

Iyengar, S., & Kinder, D. R. (2010 (1987)). *News that matters: Television and American opinion* (Updated ed.). *Chicago studies in American politics*. University of Chicago Press.

Jakobson, R. (1960). Linguistics and poetics. In T. A. Sebeok (Ed.), *MIT Press paperback series: Vol. 59. Style in language* (7th ed., pp. 350–377). MIT Press.

Jeong, M., Lin, C.-Y., & Lee, G. G. (2009). Semi-supervised speech act recognition in emails and forums. In R. C. Moore (Ed.), Proceeding *EMNLP* '09 Proceedings of the 2009 conference on empirical methods in natural language processing: Human language technology conference of the North American chapter of the association for computational linguistics (pp. 1250–1259). Association for Computational Linguistics. Retrieved from https://www.aclweb.org/anthology/D09-1130.pdf

Joachims, T. (2002). Learning to classify text using support vector machines. Zugl.: Dortmund, Univ., Diss., 2001. Kluwer international series in engineering and computer science: Vol. 668. Kluwer Academic Publishing.

Johnson, D. E., Oles, F. J., Zhang, T., & Goetz, T. (2002). A decision-tree-based symbolic rule induction system for text categorization. *IBM Systems Journal*, *41*(3), 428–437. https://doi.org/10.1147/sj.413.0428

Jones, K. S. (2004). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, *60*(5), 493–502. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.8343&rep=rep1&type=pdf

Joty, S., & Hoque, E. (2016). Speech act modeling of written asynchronous conversations with task-specific embeddings and conditional structured models. In *Proceedings of the 54th annual meeting of the association for computational linguistics* (pp. 1746–1756). Retrieved from https://www.aclweb.org/anthology/P16-1165.pdf

Kaggle: Datasets. (2015). Retrieved from https://www.kaggle.com/datasets

Kampf, Z. (2013). Mediated performatives. *Handbook of Pragmatics*. (17), 1–24. https://doi.org/10.1075/hop.17.med1

Kennedy, E. A., & Jordans, F. (2011, December 1). UN: Syria now in a civil war. *NBCNews*. Retrieved from http://www.nbcnews.com/id/45514855/ns/world_news-mideast_n_africa/t/un-syria-now-civil-war/#.Xh4Q3kdKhPa

Kepplinger, H. M., & Roth, H. (1979). Creating a crisis: German mass media and oil supply in 1973/74. *Public Opinion Quarterly*, *43*, 285–296.

Khazaei, T., Lu, X., & Mercer, R. (2017). Writing to persuade: Analysis and detection of persuasive discourse. In *iConference 2017 Proceedings* (pp.213–215). Retrieved from https://www.ideals.illinois.edu/bitstream/handle/2142/96673/1.22_346_Khazaei-Writing%20to%20Persuade.pdf?sequence=1&isAllowed=y

Khoo, A., Marom, Y., & Albrecht, D. (2006). Experiments with sentence classification. *Proceedings of the 2006 Australasian Language Technology Workshop*, 18–25. Retrieved from https://www.aclweb.org/anthology/U06-1005.pdf

Kim, S.-H., Scheufele, D. A., & Shanahan, J. (2002). Think about it this way: Attribute agenda-setting function of the press and the public's evaluation of a local issue. *Journalism & Mass Communication Quarterly*, *79*(1), 7–25. https://doi.org/10.1177/107769900207900102

Kim, S. N., Martinez, D., & Cavedon, L. (2011). Automatic classification of sentences to support evidence based medicine. In *Proceedings of the ACM fourth international workshop on data and text mining in biomedical informatics*. Retrieved from https://bmcbioinformatics.biomedcentral.com/track/pdf/10.1186/1471-2105-12-S2-S5

Kim, Y. (2014). Convolutional neural networks for sentence classification. *In Proceedings of the 2014 conference on empirical methods in natural language processing*, 1746–1752. Retrieved from https://www.aclweb.org/anthology/D14-1181.pdf

Kiousis, S., & McDevitt, M. (2008). Agenda setting in civic development: Effects of curricula and issue importance on youth voter turnout. *Communication Research*, *35*(4), 481–502. https://doi.org/10.1177%2F0093650208315978

Kiousis, S., & McCombs, M. (2004). Agenda-setting effects and attitude strength: Political figures during the 1996 presidential election. *Communication Research*, *31*(1), 36–57. https://doi.org/10.1177/0093650203260205

Klandermans, B., & Weerd, M. d. (1999). Injustice and adversarial frames in a supranational political context: Farmer's protest in the Netherlands and Spain. In D. Della Porta, H. Kriesi, & D. Rucht (Eds.), *Social movements in a globalizing world* (pp. 134–147). Macmillan.

Kohavi, R., & Provost, F. (1998). Glossary of terms. *Machine Learning*, *30*, 271–274. https://doi.org/10.1023/A:1017181826899

König, E., & Siemund, P. (2007). Speech act distinctions in grammar. In T. Shopen (Ed.), *Language typology and syntactic description* (pp. 276–324). Cambridge University Press. https://doi.org/10.1017/CBO9780511619427.005

Kosicki, G. M. (1993). Problems and opportunities in agenda-setting research. *Journal of communication*, *43*(2), 100–127. https://doi.org/10.1111/j.1460-2466.1993.tb01265.x

Krippendorff, K. (2004a). Content analysis: An introduction to its methodology. SAGE Publications, Inc.

Krippendorff, K. (2004b). Measuring the reliability of qualitative text analysis data. *Quality & Quantity*, *38*(6), 787–800. https://doi.org/10.1007/s11135-004-8107-7

Kutter, A., & Kantner, C. (2011). Corpus-based content analysis: A method for investigating news coverage on war and intervention. Retrieved from https://www.sowi.uni-stuttgart.de/dokumente/forschung/irowp/IROWP_Series_2012_1_Kutter_Kantner_Corpus-Based_Content_Analysis.pdf

Landaues, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*, 259–284. https://doi.org/10.1080/01638539809545028

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159. https://doi.org/10.2307/2529310

Lang, G. E., & Lang, K. (1983). The battle for public opinion: The president, the press, and the polls during Watergate. *American Journal of Sociology*, *90*(5), 1101–1103. https://doi.org/10.1177%2F001654928403400304

Larsson, A. O., & Moe, H. (2011). Who tweets? Tracking microblogging use in the 2010 Swedish election campaign. In *ECIS 2011 Proceedings*. 251. Retrieved from http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1250&context=ecis2011

Lavidge, R. J., & Steiner, G. A. (1961). A model for predictive measurements of advertising effectiveness. *Journal of Marketing*, *25*(6), 59. https://doi.org/10.2307/1248516

Lee, G. (2010). Who let priming out? Analysis of first- and second-level agenda setting effects on priming. *International Communication Gazette*, 72(8), 759–776. https://doi.org/10.1177/1748048510380814

Lee, N. Y. (2015). Two different motivations on agenda setting: Need for orientation and motivated reasoning. *International Journal of Public Opinion Research*, 28(4), 484–510. https://doi.org/10.1093/ijpor/edv029

Lee, S. T., & Maslog, C. C. (2005). War or peace journalism? Asian newspaper coverage of conflicts. *Journal of Communication*, *55*(2), 311–329. https://doi.org/10.1111/j.1460-2466.2005.tb02674.x

Leech, G., & Weisser, M. (2003). Generic speech act annotation for task-oriented dialogues. In D. Archer, P. Rayson, A. Wilson, & A. M. McEnery (Eds.), *Proceedings of the corpus linguistics 2003 conference* (pp. 441–446). Centre for Computer Corpus Research on Language Technical Papers, Lancaster University. Retrieved from https://www.lancaster.ac.uk/fass/doc_library/linguistics/leechg/leech_and_weisser_2003.pdf

Lewis, D. D., Yang, Y., Rose, T., & Li, F. (2004). RCV1: A new benchmark collection for text categorization research. *Journal of Machine Learning Research*. (5), 361–397. Retrieved from http://www.jmlr.org/papers/volume5/lewis04a/lewis04a.pdf

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, *4*(28), 587–604. https://doi.org/10.1111/j.1468-2958.2002.tb00826.x

Louis Charbonneau (2013, July 9). Russia: Syria rebels likely behind Aleppo chemical attack. *Reuters*. Retrieved from https://uk.reuters.com/article/uk-syria-crisis-chemical-russia/russia-syria-rebels-likely-behind-aleppo-chemical-attack-idUKBRE9680S620130709

Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. (2011). Learning word vectors for sentiment analysis. In Min-Yen Kan (Ed.), *Proceedings of the 49th annual meeting of the association for computational linguistics* (pp. 142–150). Retrieved from https://www.aclweb.org/anthology/P11-1015.pdf

Madden, T. J., Ellen, P. S., & Ajzen, I. (1992). A comparison of the theory of planned behavior and the theory of reasoned action. *Personality and Social Psychology Bulletin*, *18*(1), 3–9. https://doi.org/10.1177/0146167292181001

Manning, C. D., Raghavan, P., & Schütze, H. (2009). *Introduction to information retrieval* (Reprinted.). Cambridge University Press. Retrieved from http://reference-tree.com/book/an-introduction-to-information-retrieval?utm_source=gbv&utm_medium=referral&utm_campaign=collaboration

Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. MIT Press. Retrieved from https://www.cs.vassar.edu/~cs366/docs/Manning_Schuetze_StatisticalNLP.pdf

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., & McClosky, D. (2014). The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd annual meeting of the association for computational linguistics: System demonstrations* (pp. 55–60). Retrieved from https://www.aclweb.org/anthology/P14-5010.pdf

Marion, R. (2004). Developing research questions. In *The whole art of deduction. Research skills for new scientists*. Retrieved from https://smithcreekstudios.com/wad/wad_home.htm

Matthes, J., & Kohring, M. (2008). The content analysis of media frames: Toward improving reliability and validity. *Journal of Communication*, *58*(2), 258–279. https://doi.org/10.1111/j.1460-2466.2008.00384.x

Mazzetti, M., Schmitt, E., & Banco, E. (2013, July 15). No quick impact in U.S. arms plan for Syria rebels. *New York Times*. Retrieved from http://www.nytimes.com/2013/07/15/world/middleeast/no-quick-impact-in-us-arms-plan-for-syria-rebels.html

McCallum, A., & Nigam, K. (1998). A comparison of event models for naïve Bayes text classification. In *Proceedings AAAI-98 Workshop on Learning for Text Categorization*, 41–48. Retrieved from https://www.cs.cmu.edu/~knigam/papers/multinomial-aaaiws98.pdf

McCombs, M., Llamas, J. P., Lopez-Escobar, E., & Rey, F. (1998). Candidate images in Spanish elections: Second-level agenda-setting effects. *Journalism & Mass Communication Quarterly*, *74*(4), 703–717. https://doi.org/10.1177%2F107769909707400404

McCombs, M., & Evatt, D. (1995). Issues and attributes: Exploring a new dimension in agenda setting. *Communication Sociedad*, 8, 7–32.

McCombs, M. (2004). Setting the agenda: The mass media and public opinion. Polity.

McCombs, M. E. (2005). A Look at agenda-setting: Past, present and future. *Journalism Studies*, *6*(4), 543–557. https://doi.org/10.1080/14616700500250438

McCombs, M. E., & Estrada, G. (1997). The news media and the pictures in our heads. In S. Iyengar & B. Reeves (Eds.), *Do the media govern?* (pp. 237–247). SAGE Publications.

McCombs, M. E., Lopez-Escobar, E., & Llamas, J. P. (2000). Setting the agenda of attributes in the 1996 Spanish general election. *Journal of Communication*, *50*(2), 77–92. https://doi.org/10.1111/j.1460-2466.2000.tb02842.x

McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly*, *36*(2), 176. https://doi.org/10.1086/267990

McCombs, M. E., & Shaw, D. L. (1993). The evolution of agenda-setting research: Twenty-five years in the marketplace of ideas. *Journal of Communication*, *43*(2), 58–67. https://doi.org/10.1111/j.1460-2466.1993.tb01262.x

McCombs, M. E., & Stroud, N. J. (2014). Psychology of agenda-setting effects: Mapping the paths of information processing. *Review of Communication Research*, *2*(1), 68–93. Retrieved from https://www.ssoar.info/ssoar/bitstream/handle/document/37343/ssoar-rcr-2014-1-mccombs_et_al-Psychology_of_agenda-setting_effects_mapping.pdf?sequence=1&isAllowed=y&lnkname=ssoar-rcr-2014-1-mccombs_et_al-Psychology_of_agenda-setting_effects_mapping.pdf

McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field*. SAGE Publications, Inc. http://dx.doi.org/10.4135/9781452225913

McGowan, M. K., Tam, S. S., & Hall, M. (2009). On indirect speech acts and linguistic communication: A response to Bertolet. *Philosophy*, *84*(04), 495. https://doi.org/10.1017/S0031819109990088

McGuire, W. J. (1986). The myth of massive media impact: Savagings and salvagings. In G. Comstock (Ed.), *Public communication and behavior* (pp. 173–257). Acad. Press. https://doi.org/10.1016/B978-0-12-531956-0.50007-X

McKnight, L., & Srinivasan, P. (2003). Categorization of sentence types in medical abstracts. In *Proceedings of AMIA annual symposium*, 44–444. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1479904/

Meacham, J., & Leiman, B. (1982). Remembering to perform future actions. In U. Neisser (Ed.), *Memory observed: Remembering in natural contexts* (pp. 327–336). Freeman.

Meijers, A. (2007). Collective Speech Acts. In S. L. Tsohatzidis & S. L. Tsochatzidēs (Eds.), *Theory and decision library: Vol. 41. Intentional acts and institutional facts: Essays on John Searle's social ontology* (pp. 93–110). Springer. https://doi.org/10.1007/978-1-4020-6104-2_4

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient estimation of word representations in vector space*. Retrieved from http://arxiv.org/pdf/1301.3781

Minsky, M. L., & Papert, S. A. (1990). *Perceptrons: An introduction to computational geometry* (5th, expanded ed.). MIT Press.

Moon, S. J. (2008). Agenda-setting effects as a mediator of media use and civic engagement: From what the public thinks about to what the public does (Doctoral dissertation). The University of Texas at Austin.

Moon, S. J. (2013). Attention, attitude, and behavior: Second-level agenda-setting effects as a mediator of media use and political participation. *Communication Research*, *40*(5), 698–719. https://doi.org/10.1177/0093650211423021

Moon, S. J. (2014). From what the public thinks about to what the public does: Agenda setting effects as a mediator of media use and civic engagement. In T. J. Johnson (Ed.), *Agenda setting in a 2.0 world: New agendas in communication; a tribute to Maxwell McCombs* (pp. 158–189). Routledge.

Morris, P. E. (1992). Prospective memory: Remembering to do things. In M. Gruneberg & P. E. Morris (Eds.), *Aspects of memory. Vol.1: The practical aspects* (pp. 196–222). Routledge.

Murphy, J. (2014). (Im)politeness during Prime Minister's questions in the U.K. Parliament. *Pragmatics and Society*, *5*(1), 76–104. https://doi.org/10.1075/ps.5.1.04mur

Nagarajan, M., Sheth, A., Aguilera, M., Keeton, K., Merchant, A., & Uysal, M. (2007). Altering document term vectors for classification – ontologies as expectations of co-occurrence. In C. Williamson, M. E. Zurko, P. Patel-Schneider, & P. Shenoy (Eds.), *Proceedings of the 16th international conference on World Wide Web* (pp. 1225–1226). https://doi.org/10.1145/1242572.1242778

Nelson, T. E., Clawson, R. A., & Oxley, Z. M. (1997). Media framing of a civil liberties conflict and its effect on tolerance. *The American Political Science Review*, *91*(3), 567–583. Retrieved from https://www.uvm.edu/~dguber/POLS234/articles/nelson.pdf

Neuendorf, K. A. (2002). *The content analysis guidebook*. SAGE Publications, Inc. Retrieved from https://academic.csuohio.edu/kneuendorf/SkalskiVitae/SkalskiNeuendorfCajigas17.pdf

Nigam, K., Mccallum, A. K., Thrun, S., & Mitchell, T. (2000). Text classification from labeled and unlabeled documents using EM. *Machine Learning*, *39*(2/3), 103–134. https://doi.org/10.1023/A:1007692713085

Obama warns Syria not to cross 'red line' (2012, August 21). *CNN*. Retrieved from https://edition.cnn.com/2012/08/20/world/meast/syria-unrest/

Obeng, S. G. (1997). Language and politics: Indirectness in political discourse. *Discourse & Society*, *8*(1), 49–83. https://doi.org/10.1177/0957926597008001004

Onions, C. T. (Ed.). (1996). *The Oxford dictionary of English etymology* (Repr). Oxford: Oxford Univ. Press.

Oya, T., & Carenini, G. (2014). Extractive summarization and dialogue act modeling on email threads: An integrated probabilistic approach. In *Proceedings of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)* (pp. 133–140). Association for Computational Linguistics. Retrieved from https://www.aclweb.org/anthology/W14-4318.pdf

Pawar, P. Y., & Gawande, S. H. (2012). A comparative study on different types of approaches to text categorization. *International Journal of Machine Learning and Computing*, *2*(4), 423–426. Retrieved from http://www.ijmlc.org/papers/158-C01020-R001.pdf

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Édouard Duchesnay. (2011). Scikit-learn: Machine learning in Python. *JMLR*, *12*, 2825–2830. Retrieved from http://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf

Phan, X.-H., Nguyen, L.-M., & Horiguchi, S. (2008). Learning to classify short and sparse text & web with hidden topics from large-scale data collections. In J. Huai, R. Chen, H.-W. Hon, Y. Liu, W.-Y. Ma, A. Tomkins, & X. Zhang (Eds.), *Proceeding of the 17th international conference on World Wide Web* (pp. 91–100). https://doi.org/10.1145/1367497.1367510

Phan, X.-H., Nguyen, C.-T., Le, D.-T., Nguyen, L.-M., Horiguchi, S., & Ha, Q.-T. (2011). A hidden topic-based framework toward building applications with short web documents. *IEEE transactions on knowledge and data engineering*, *23*(7), 961–976. https://doi/ 10.1109/TKDE.2010.27

Pinto, D., Jiménez-Salazar, H., & Rosso, P. (2006). Clustering abstracts of scientific texts using the transition point technique. In *Computational linguistics and intelligent text processing* (pp. 536–546). Retrieved from https://www.cs.buap.mx/~dpinto/research/PaperUPV6/Pinto.pdf

Portner, P. (2004). The semantics of imperatives within a theory of clause types. In R. Young (Ed.), *Semantics and linguistic theory (SALT)* (pp. 232–252). Retrieved from https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/2907/2647

Pu, Q., & Yang, G.-W. (2006). Short-text classification based on ica and lsa. In J. Wang, Z. Yi, J. M. Zurada, B.-L. Lu, & H. Yin (Eds.), *Advances in neural networks – ISNN 2006, volume 3972 of lecture notes in computer science* (pp. 265–270). Springer. https://doi.org/10.1007/11760023_39

Qadir, A., & Riloff, E. (2011). Classifying sentences as speech acts in message board posts. *Proceedings of the 2011 conference on empirical methods in natural language processing*, 748–759. Retrieved from https://www.aclweb.org/anthology/D11-1069.pdf

Ramírez-de-la-Rosa, G., Montes-y-Gómez, M., Solorio, T., & Villaseñor-Pineda, L. (2013). A document is known by the company it keeps: Neighborhood consensus for short text categorization. *Lang Resources & Evaluation*, *47*(1), 127–149. https://doi.org/10.1007/s10579-012-9192-1

Reese, S. (1991). Setting the media's agenda: A power balance perspective. In J. Anderson (Ed.), *Communication yearbook 14* (pp. 309–340). SAGE Publications, Inc.

Reidsma, D., & Carletta, J. (2008). Reliability measurement without limits. *Computational linguistics*, *34*(3), 319–326. https://doi.org/10.1162/coli.2008.34.3.319

Reitermanova, Z. (2010). Data Splitting. In J. Šafránková (Ed.), 19th annual conference of doctoral students, *WDS*'10 'Week of Doctoral Students 2010', Charles University, faculty of mathematics and physics, Prague, Czech Republic, June 1, 2010 to June 4, 2010: [proceedings of contributed papers] (pp. 31–36). Matfyzpress.

Rennie, J. D. M., Shih, L., Teevan, J., & Karger, D. R. (2003). Tackling the poor assumptions of naïve Bayes text classifiers. In *Proceedings of the twentieth international conference on machine learning* (pp. 616–623). Retrieved from https://people.csail.mit.edu/jrennie/papers/icml03-nb.pdf

Revathi, T., Ramya, L. V., Tanuja, M., Pavani, S., & Swathi, M. (2012). Sentence level semantic classification of online product reviews of mixed opinions using naïve Bayes classifier. *International Journal of Engineering Trends and Technology (IJETT)*, *3*(2), 127–132. Retrieved from https://pdfs.semanticscholar.org/5d88/bc396df30615bed29716c653be09e2472f5e.pdf?_ga=2.233677342.1406669603.1578836404-1285514259.1574199049

Roberts, C. (2017). Speech acts in discourse context. In D. Fogal, D. Harris, & M. Moss (Eds.), *New work on speech acts.* Oxford University Press. Retrieved from https://www.asc.ohio-state.edu/roberts.21/Roberts.SpeechActs.pdf

Roberts, M. S. (1992). Predicting voting behavior via the agenda-setting tradition. *Journalism Quarterly*, *69*(4), 878–892. https://doi.org/10.1177%2F107769909206900408

Rogers, E. M., & Dearing, J. W. (1988). Agenda-setting research: Where has it been, where is it going? In *Communication yearbook 11* (pp. 555–594). https://doi.org/10.1080/23808985.1988.11678708

Rogers, E. M., Dearing, J. W., & Chang, S. (1991). *AIDS in the 1980s: The agenda-setting process for a public issue*. Association for Education in Journalism and Mass Communication.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386–408. https://doi.org/10.1037/h0042519

Rosso, P., Errecalde, M., & Pinto, D. (2013). Analysis of short texts on the Web: Introduction to special issue. *Language Resources and Evaluation*, *47*(1), 123–126. https://doi.org/10.1007/s10579-013-9220-9

Ruiz, M. E., & Srinivasan, P. (2002). Hierarchical text categorization using neural networks. *Information Retrieval*, *5*(1), 87–118. https://doi.org/10.1023/A:1012782908347

Sadock, J. M. (1974). Toward a linguistic theory of speech acts. Academic Press.

Sadock, J. M. (1994). Toward a grammatically realistic typology of speech acts. In S. L. Tsohatzidis (Ed.), *Foundations of speech act theory: Philosophical and linguistic perspectives* (pp. 393–406). Routledge.

Sahlgren, M., & Cöster, R. (2004). Using bag-of-concepts to improve the performance of support vector machines in text categorization. In *Proceedings of the 20th international conference on Computational Linguistics* (pp. 1–7). https://doi.org/10.3115/1220355.1220425

Saif, H., Fernández, M., He, Y., & Alani, H. (2014). On stopwords, filtering and data sparsity for sentiment analysis of Twitter. In N. Calzolari (Ed.), *Proceedings of the ninth international conference on language resources and evaluation (LREC 2014)* (pp. 810–817). Retrieved from http://www.lrec-conf.org/proceedings/lrec2014/pdf/292_Paper.pdf

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, *24*(5), 513–523. https://doi.org/10.1016/0306-4573(88)90021-0

Salton, G., & McGill, M. J. (1986). *Introduction to modern information retrieval*. McGraw-Hill, Inc.

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, *44*(1.2), 210–229.

Sanfilippo, A., Franklin, L., Tratz, S., Danielson, G., Mileson, N., Riensche, R., & McGrath, L. (2008). Automating frame analysis. In H. Liu, J. J. Salerno, & M. J. Young (Eds.), *Social computing, behavioral modeling, and prediction* (pp. 239–249). Springer.

Santorini, B. (1990). Part-of-speech tagging guidelines for the Penn Treebank project. In *Technical report MS-CIS-90-47*. Department of Computer and Information Science, University of Pennsylvania. Retrieved from https://catalog.ldc.upenn.edu/docs/LDC99T42/tagguid1.pdf

Saussure, F. d. (1916 (1959)). *Course in general linguistics*. Philosophical Library.

Sbisà, M. (1995). Speech act theory. In J. Verschueren, J.-O. Östman, & J. Blommaert (Eds.), *Handbook of pragmatics* (pp. 495–506). John Benjamins Publishing Company.

Sbisà, M. (2002). Speech acts in context. *Language & Communication*, *22*(4), 421–436. https://doi.org/10.1016/S0271-5309(02)00018-6

Scharkow, M. (2012). *Automatische Inhaltsanalyse und maschinelles Lernen* (Doctoral dissertation, Universität der Künste, Berlin). Retrieved from https://opus4.kobv.de/opus4-udk/frontdoor/deliver/index/docId/28/file/dissertation_scharkow_final_udk.pdf

Scharkow, M. (2013). Thematic content analysis using supervised machine learning: An empirical evaluation using German online news. *Quality & Quantity*, *47*(2), 761–773. https://doi.org/10.1007/s10579-012-9192-1

Scheufele, D. A. (1999). Framing as a theory of media effects. *Journal of Communication*, *49*(1), 103–122. https://doi.org/10.1111/j.1460-2466.1999.tb02784.x

Scheufele, D. A. (2000). Agenda-setting, priming, and framing revisited: Another look at cognitive effects of political communication. *Mass Communication and Society*, *3*(2-3), 297–316. https://doi.org/10.1207/S15327825MCS0323_07

Schonhofen, P. (2006). Identifying document topics using the Wikipedia category network. In *2006 IEEE/WIC/ACM international conference on Web intelligence (WI 2006 Main Conference Proceedings)(WI'06)* (pp. 456–462). https://doi.org/10.1109/WI.2006.92

Scott, W. A. (1955). Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quarterly*, *19*(3), 321–325.
Searle, J. R. (1975a). A taxonomy of illocutionary acts. In K. Günderson (Ed.), *Language, mind, and knowledge: (Minneapolis studies in the philosophy of science, vol. 7)* (pp. 344–369). University of Minneapolis Press.
Searle, J. R. (1975b). Indirect speech acts. *Syntax and Semantics, 3: Speech Acts*, *3*, 59–82.
Searle, J. R. (1976). A classification of illocutionary acts. *Language in Society*, *5*(1), 1–23. Retrieved from http://www.jstor.org/stable/4166848?origin=JSTOR-pdf
Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys*, *34*(1), 1–47. https://doi.org/10.1145/505282.505283
Segaran, T. (2007). *Programming collective intelligence: Building smart web 2.0 applications* (1. ed.). *Safari tech books online*. O'Reilly. Retrieved from http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10758612
Severin, W. J., & Tankard, J. W. (2001). Communication theories: Origins, methods, and uses in the mass media (5. ed.). Longman.
Shafiq, J., Carenini, G., & Lin, C.-Y. (2011). Unsupervised modeling of dialog acts in asynchronous conversations. In T. Walsh (Ed.), *Proceedings of the twenty-second international joint conference on artificial intelligence, IJCAI-11: Barcelona, Catalonia, Spain, 16–22 July 2011* (p. 1807/1813). AAAI Press. Retrieved from https://pdfs.semanticscholar.org/3cdd/db616be3f672a9fa0e4b82a8717b2b513d92.pdf
Shoemaker, P. J. (Ed.). (1989). Communication campaigns about drugs: Government, media, and the public. Communication textbook series. Erlbaum.
Siddique, H. (2013, September 16). Doctors to warn that Syria's healthcare system is at 'breaking point'. *The Guardian*. Retrieved from https://www.theguardian.com/world/2013/sep/16/doctors-warn-syria-healthcare-breaking-point
Silva, C., & Ribeiro, B. (2003). The importance of stop word removal on recall values in text categorization. In Proceedings of the international joint conference on neural networks 2003: Doubletree Hotel,

Jantzen Beachm Portland, Oregon, July 20 - 24, 2003 (pp. 1661–1666). IEEE Service Center.

Silveira, N., Dozat, T., Marneffe de, M.-C., Bowman, S. R., Connor, M., Bauer, J., & Manning, C. D. (2014). A gold standard dependency corpus for English. In N. Calzolari (Ed.), *Proceedings of the ninth international conference on language resources and evaluation (LREC 2014), Reykjavik, Iceland, May 26-31, 2014*. Retrieved from http://www.lrec-conf.org/proceedings/lrec2014/pdf/1089_Paper.pdf

Smith, B. (1990). Towards a history of speech act theory. In A. Burkhardt (Ed.), *Speech acts, meanings and intentions. Critical approaches to the philosophy of John R. Searle* (pp. 29–61). de Gruyter.

Snow, D. A., & Benford, R. D. (1988). Ideology, frame resonance, and participant mobilization. *International Social Movement Research*, *1*(1), 197–217.

Snow, D. A., & Benford, R. D. (1992). Master frames and cycles of protest. In A. D. Morris & C. M. Mueller (Eds.), *Frontiers in social movement theory* (pp. 133–155). Yale University Press.

Snyder, A. C. (2017). *Setting the Agenda for Global Peace*. Taylor and Francis.

Song, G., Ye, Y., Du, X., Huang, X., & Bie, S. (2014). Short text classification: A survey. *Journal of Multimedia*, *9*(5), 635–643. https://doi.org/10.4304/jmm.9.5.635-643

Soucy, P., & Mineau, G. W. (2001). A simple *k*NN algorithm for text categorization. In N. Cercone (Ed.), *Proceedings / 2001 IEEE international conference on data mining: 29 November – 2 December 2001, San Jose, California* (pp. 647–648). IEEE Computer Society. https://doi.org/10.1109/ICDM.2001.989592

Sriram, B., Fuhry, D., Demir, E., Ferhatosmanoglu, H., & Demirbas, M. (2010). Short text classification in twitter to improve information filtering. In F. Crestani, S. Marchand-Maillet, H.-H. Chen, E. N. Efthimiadis, & J. Savoy (Eds.), *Proceeding of the 33rd international ACM SIGIR conference* (pp. 841–850). https://doi.org/10.1145/1835449.1835643

Stalpouskaya, K., & Baden, C. (2015). To do or not to do: The role of agendas for action in analyzing news coverage of violent conflict. In Min-Yen Kan (Ed.), *Proceedings of the 53rd annual meeting of the* *association for computational linguistics and the 7th international joint conference on natural language processing: Proceedings of the first workshop on computing newsstorylines (CNewsStory 2015)* (pp. 21–30). Retrieved from https://www.aclweb.org/anthology/W15-4504.pdf

Steinberg, M. W. (1999). The talk and back talk of collective action: A dialogic analysis of repertoires of discourse among nineteenth century English cotton spinners 1. *American Journal of Sociology*, *105*(3), 736–780. https://doi.org/10.1086/210359

Stevenson, A., & Pearsall, J. (Eds.). (2010). *Oxford dictionary of English* (3. ed.). Oxford University Press.

Stieglitz, S., & Dang-Xuan, L. (2013). Social media and political communication: A social media analytics framework. *Social Network Analysis and Mining*, *3*(4), 1277–1291. https://doi.org/10.1007/s13278-012-0079-3

Stolcke, A., Ries, K., Coccaro, N., Shriberg, E., Bates, R., Jurafsky, D., ... Meteer, M. (2000). Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics*, *26*(3), 339–373. Retrieved from https://www.aclweb.org/anthology/J00-3003.pdf

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society*, *36*(2), 111–147.

Stroud, N. J., & Kenski, K. (2007). From agenda setting to refusal setting. *Public Opinion Quarterly*, *71*(4), 539–559. https://doi.org/10.1093/poq/nfm042

Sutherland, M., & Galloway, J. (1981). Role of advertising: Persuasion or agenda setting? *Journal of Advertising Research*, *21*(5), 25–29.

Syria in civil war, says UN official Herve Ladsous (2012, June 12). *BBC*. Retrieved from https://www.bbc.com/news/world-middle-east-18417952

Takeshita, T. (1997). Exploring the media's roles in defining reality: From issue-agenda setting to attribute-agenda setting. In M. E. McCombs, D. L. Shaw, & D. H. Weaver (Eds.), *Communication and democracy: Exploring the intellectual frontiers in agenda-setting theory* (pp. 15–27). Lawrence Erlbaum Associates.

Takeshita, T. (2006). Current critical problems in agenda-setting research. *International Journal of Public Opinion Research*, *18*(3), 275–296. https://doi.org/10.1093/ijpor/edh104

Tavafi, M., Mehdad, Y., Joty, S., Carenini, G., & Ng, R. (2013). Dialogue act recognition in synchronous and asynchronous conversations. In *Proceedings of the SIGDIAL 2013 conference* (pp. 117–121). Retrieved from https://www.aclweb.org/anthology/W13-4017.pdf

Tenenboim-Weinblatt, K. (2013). Bridging collective memories and public agendas: Toward a theory of mediated prospective memory. *Communication Theory*, *23*(2), 91–111. https://doi.org/10.1111/comt.12006

Tesnière, L. (1976 [1959]). Éléments *de syntaxe structurale [Elements of Structural Syntax]* (2. éd. rev. et corr., 3. tirage). Klincksieck.

Tewksbury, D., & Scheufele, D. (2009). News framing theory and research. In J. Bryant & M. B. Oliver (Eds.), *Media effects: Advances in theory and research* (pp. 17–33). Routledge.

The Merriam Webster dictionary of synonyms and antonyms. (1992). A Merriam-Webster. Merriam-Webster.

Ting, S. L., Ip, W. H., & Tsang, A. H. (2011). Is naïve Bayes a good classifier for document classification? *International Journal of Software Engineering and Its Applications*, *5*(3), 37–46.

Trstenjak, B., Mikac, S., & Donko, D. (2014). *k*NN with TF-IDF based framework for text categorization. *Procedia Engineering*, *69*, 1356–1364. https://doi.org/10.1016/j.proeng.2014.03.129

Tsohatzidis, S. L. (Ed.). (1994). Foundations of speech act theory: Philosophical and linguistic perspectives. Routledge.

Tsur, O., Calacci, D., & Lazer, D. (2015). A frame of mind: Using statistical models for detection of framing and agenda setting campaigns. In Min-Yen Kan (Ed.), *Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing* (pp. 1629–1638). Retrieved from https://www.aclweb.org/anthology/P15-1157.pdf

UNHRC. (2014). *Independent International Commission of Inquiry on the Syrian Arab Republic*. Retrieved from https://www.ohchr.org/EN/HRBodies/HRC/IICISyria/Pages/Documentation.aspx

Van Atteveldt, W. (2008). Semantic network analysis: Techniques for extracting, representing and querying media content. *SIKS* dissertation series: no. 2008-30. BookSurge. Retrieved from http://vanatteveldt.com/uploads/vanatteveldt_semanticnetworkanalysis.pdf

Van Atteveldt, W., Sheafer, T., Shenhav, S. R., & Fogel-Dror, Y. (2017). Clause analysis: Using syntactic information to automatically extract source, subject, and predicate from texts with an application to the 2008–2009 Gaza War. *Political Analysis*, *25*(02), 207–222. https://doi.org/10.1017/pan.2016.12

Van der Pas, D. (2014). *From the press to politics and back: When do media set the political agenda and when do parties set the media agenda?* [S.l], Amsterdam: [s.n.]; Universiteit van Amsterdam [Host]. Retrieved from https://dare.uva.nl/search?identifier=ed0f62ec-7d45-44c5-b664-8f2c2b603a0f

Van Dijk, T. A. (1977). Context and cognition: Knowledge frames and speech act comprehension. *Journal of Pragmatics*, *1*(3), 211–231. https://doi.org/10.1016/0378-2166(77)90035-2

Van Rijsbergen, C. J. (1979). *Information retrieval* (2. ed., repr). Butterworth.

Vargo, C. J., Guo, L., McCombs, M., & Shaw, D. L. (2014). Network issue agendas on Twitter during the 2012 U.S. presidential election. *Journal of Communication*, *64*(2), 296–316. https://doi.org/10.1111/jcom.12089

Vateekul, P., & Kubat, M. (2009). Fast induction of multiple decision trees in text categorization from large scale, imbalanced, and multi-label data. In Y. Saygin (Ed.), IEEE *international conference on data mining workshops, 2009:* ICDMW *'09; 6 Dec. 2009, Miami, Florida, USA* (pp. 320–325). IEEE. https://doi.org/10.1109/ICDMW.2009.94

Verschueren, J. (1980). *On speech act verbs* (I:4). John Benjamins Publishing Company.

Walters-Williams, J., & Li, Y. (2010). Comparative study of distance functions for nearest neighbors. In K. Elleithy (Ed.), *Advanced Techniques in Computing Sciences and Software Engineering*. Springer. https://doi.org/10.1007/978-90-481-3660-5_14

Wang, J., Zhou, Y., Li, L., Hu, B., & Hu, X. (2009). Improving short text clustering performance with keyword expansion. In J. Kacprzyk, T. Huang, Y. Shen, H. Wang, & Z. Zeng (Eds.), *Advances in intelligent and soft computing: Vol. 56. The Sixth International Symposium on Neural Networks (*ISNN *2009)* (pp. 291–298). Springer. https://doi.org/10.1007/978-3-642-01216-7_31

Weaver, D. H. (1977). Political issues and voter need for orientation. In D. L. Shaw & M. E. McCombs (Eds.), *The emergence of American political issues: The agenda-setting function of the press* (pp. 107–119). West Publishing Co.

Weaver, D. H. (1991). Issue salience and public opinion: Are there consequences of agenda setting. *International Journal of Public Opinion Research*, *3*(1), 53–68. https://doi.org/10.1093/ijpor/3.1.53

Weaver, D. H. (2007). Thoughts on agenda setting, framing, and priming. *Journal of Communication*, *57*(1), 142–147. https://doi.org/10.1111/j.1460-2466.2006.00333.x

Weaver, D. H., & Choi, J. (2017). The media agenda: Who (or what) sets it? In K. Kenski, K. H. Jamieson, D. H. Weaver, & J. Choi (Eds.), *The Oxford handbook of political communication* (Vol. 1). Oxford University Press. https://dx.doi.org/10.1093/oxfordhb/9780199793471.013.37

Wierzbicka, A. (1987). English speech act verbs: A semantic dictionary. Academic Press.

Wilson, J. (1990). Politically speaking: The pragmatic analysis of political language. Language in society: Vol. 16. Blackwell.

Witten, I. H., Moffat, A., & Bell, T. C. (1999). Managing gigabytes: Compressing and indexing documents and images (2. ed., [Nachdr.]). The Morgan Kaufmann series in multimedia information and systems. Morgan Kaufmann.

Wood, A., Rodeghero, P., Armaly, A., & McMillan, C. (2018, November). *Detecting speech act types in developer question/answer conversations during bug repair.* Paper presented at the ESEC/FSE.

Wu, H. C., Luk, R. W. P., Wong, K. F., & Kwok, K. L. (2008). Interpreting TF-IDF term weights as making relevance decisions. ACM *Transactions on information systems*, *26*(3), 1–37. https://doi.org/10.1145/1361684.1361686

Xu, W., Sun, H., Deng, C., & Tan, Y. (2017). Variational autoencoder for semi-supervised text classification. In *Proceedings of the thirty-first AAAI conference on artificial intelligence (*AAAI*-17)* (pp. 3358–3364). Retrieved from https://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14299/14261

Yan, T., & Wang, X.-W. (2010). Feature extension for short text. In Y. Zou (Ed.), Proceedings of the third international symposium computer science and computational technology, (ISCSCT 2010): 14 –15 Aug. 2010, Jiaozuo, China (pp. 338–341). Academy Publ.

Zaman, A. N. K., Matsakis, P., & Brown, C. (2011). Evaluation of stop word lists in text retrieval using latent semantic indexing. In *2011 sixth international conference on digital information management (ICDIM)* (pp. 133–136). https://doi.org/10.1109/ICDIM.2011.6093315

Zelikovitz, S., & Hirsh, H. (2000). Improving short text classification using unlabeled background knowledge to assess document similarity. In *Proceedings of the seventeenth international conference on machine learning* (pp. 1183–1190). Retrieved from https://pdfs.semanticscholar.org/e570/e51993722e194d8465149216d3d7351b824c.pdf

Zelikovitz, S., & Marquez, F. (2005). Transdutive learning for short-text classification problems using latent semantic indexing. *International Journal of Pattern Recognition and Artificial Intelligence*, *19*(02), 143–163. https://doi.org/10.1142/S0218001405003971

Zhang, H. (2004). The Optimality of naïve Bayes. *In Proceedings of 17th international FLAIRS conference*. Retrieved from http://www.cs.unb.ca/profs/hzhang/publications/FLAIRS04ZhangH.pdf

Zhu, X., Vondrick, C., Fowlkes, C. C., & Ramanan, D. (2015). Do we need more training data? *International Journal of Computer Vision*, 1–17. https://doi.org/10.1007/s11263-015-0812-2

Appendix 1

Table 8: Confusion matrix for 100 sentences used to measure inter-coder agreement

Category/ Sentence	De-escalation	Escalation	Support	Punishment	Ignorance	General	Other	Not agendas	Multiclass	Not doing
1	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0
3	0.0	0.0	1.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
6	0.0	0.0	0.0	0.0	1.0	0.0	2.0	0.0	0.0	0.0
7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
8	1.0	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0
9	0.0	0.0	0.0	0.0	0.0	0.0	1.0	2.0	0.0	0.0
10	0.0	0.0	1.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0
11	0.0	0.0	0.0	2.0	0.0	0.0	0.0	1.0	0.0	0.0
12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	2.0	0.0
13	0.0	0.0	1.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0
14	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
15	0.0	0.0	2.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0
17	0.0	0.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
18	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	1.0
19	2.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0
20	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0	0.0
21	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
22	0.0	0.0	1.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0
23	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0

Category/ Sentence	De-escalation	Escalation	Support	Punishment	Ignorance	General	Other	Not agendas	Multiclass	Not doing
24	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0
25	0.0	0.0	0.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0
26	1.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0
27	0.0	0.0	0.0	0.0	0.0	2.0	1.0	0.0	0.0	0.0
28	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	2.0	0.0
29	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
31	1.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0
32	0.0	0.0	0.0	0.0	0.0	2.0	0.0	1.0	0.0	0.0
33	2.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
34	0.0	0.0	0.0	2.0	1.0	0.0	0.0	0.0	0.0	0.0
35	0.0	1.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0
36	0.0	1.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0
37	0.0	0.0	0.0	2.0	1.0	0.0	0.0	0.0	0.0	0.0
38	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
39	0.0	0.0	2.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0
40	1.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0
41	0.0	0.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
42	0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0
43	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	1.0
44	0.0	0.0	2.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
45	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0	0.0	0.0
46	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0	0.0
47	0.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
48	0.0	0.0	0.0	0.0	1.0	0.0	0.0	1.0	0.0	1.0
49	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	2.0	0.0

Category/ Sentence	De-escalation	Escalation	Support	Punishment	Ignorance	General	Other	Not agendas	Multiclass	Not doing
50	0.0	0.0	0.0	2.0	0.0	0.0	0.0	1.0	0.0	0.0
51	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	1.0
52	0.0	0.0	0.0	0.0	1.0	0.0	1.0	1.0	0.0	0.0
53	0.0	0.0	0.0	2.0	0.0	0.0	1.0	0.0	0.0	0.0
54	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
55	2.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
56	0.0	0.0	0.0	0.0	0.0	1.0	0.0	2.0	0.0	0.0
57	1.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
58	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0	0.0
59	0.0	0.0	0.0	0.0	0.0	0.0	1.0	2.0	0.0	0.0
60	1.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0
61	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	1.0	0.0
62	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
63	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
64	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
65	0.0	0.0	1.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0
66	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
67	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
68	0.0	0.0	1.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0
69	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
70	0.0	0.0	0.0	0.0	0.0	1.0	0.0	2.0	0.0	0.0
71	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
72	0.0	0.0	0.0	0.0	0.0	2.0	0.0	1.0	0.0	0.0
73	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
74	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0
75	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0

Category/ Sentence	De-escalation	Escalation	Support	Punishment	Ignorance	General	Other	Not agendas	Multiclass	Not doing
76	0.0	0.0	0.0	0.0	1.0	0.0	0.0	2.0	0.0	0.0
77	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
78	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	1.0
79	0.0	0.0	0.0	0.0	2.0	0.0	0.0	1.0	0.0	0.0
80	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0
81	0.0	0.0	0.0	0.0	2.0	0.0	1.0	0.0	0.0	0.0
82	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	1.0
83	0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	1.0
84	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	1.0	1.0
85	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0
86	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
87	0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0
88	0.0	0.0	2.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0
89	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	2.0
90	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	2.0	0.0
91	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	1.0
92	1.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0
93	0.0	0.0	0.0	1.0	0.0	0.0	2.0	0.0	0.0	0.0
94	0.0	0.0	0.0	0.0	0.0	1.0	0.0	2.0	0.0	0.0
95	1.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0
96	0.0	0.0	0.0	0.0	1.0	0.0	0.0	2.0	0.0	0.0
97	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0
98	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	1.0	0.0
99	0.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0
100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	2.0	0.0

Appendix 2

Speech act hint verbs	demand, request, command, urge, implore, plead, insist, hope, pray, impose, forbid, instruct, refuse, pursue, threaten, require, order, encourage, discourage, warn, strive, provoke, welcome, applause, incite, threat, call, ask, have, ultimate, stipulate, ultimatum, stipulation, insistence, dictate, pressure, clamour, clamour, menace, intimidate, browbeat, bully, pressurize, terrorize, frighten, scare, alarm, oblige, tell, want, suppose, promise, swear, vow, offer, suggest, plan, intent, pledge, guarantee, engage, coerce
Hint adjectives	pointless, senseless, futile, hopeless, fruitless, useless, needless, in vain, unavailing, necessary, unnecessary, unacceptable, imperative, obligatory, requisite, compulsory, mandatory
Verbs meaning 'to plan'	plan, go
Verbs meaning 'stand'	stand, withstand, endure, tolerate, put up
Hint nouns	priority, answer, way, solution, approach, strategy, action, measure, step
Adjectives meaning 'vital'	vital, important, significant, essential, substantial, principal, salient
Verbs meaning 'blame'	blame, condemn, deplore, decry, denounce
State verbs	be, feel
Verbs meaning 'let'	let, do not (don't)
Nouns meaning 'promise'	promise, word, assurance, vow, pledge, guarantee, oath
Verbs meaning 'give'	give, make, hold

Table 9: Lists of hint words used for linguistic features

Appendix 3

The results of different steps of hierarchical classification (see Figure 9) for five classifiers: Naïve Bayes, SVM, *k*NN, DT and MLP.

	TF-IDF for NB			TF-IDF for SVM			TF-IDF for *k*NN			TF-IDF for DT			TF-IDF for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
Cooperative treatment	0.5	0.01	0.03	0.61	0.4	0.48	0.49	0.37	0.42	0.41	**0.51**	0.46	**0.72**	0.23	0.34
Restrictive treatment	**1**	0.02	0.05	0.65	0.3	0.41	0.38	0.26	0.31	0.47	**0.43**	**0.45**	0.92	0.14	0.24
Other	0.65	**0.99**	0.78	0.73	0.89	**0.8**	0.72	0.82	0.77	0.74	0.67	0.7	0.69	0.96	**0.8**
Average	0.63	0.64	0.52	0.69	**0.71**	0.68	0.63	0.65	0.63	0.63	0.61	0.61	**0.71**	0.69	0.63
	TF-IDF + data origin for NB			TF-IDF + data origin for SVM			TF-IDF + data origin for *k*NN			TF-IDF + data origin for DT			TF-IDF + data origin for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**Precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
Cooperative treatment	0.5	0	0.01	0.59	0.44	**0.51**	0.54	0.41	0.47	0.38	0.47	0.42	0.58	0.42	0.49
Restrictive treatment	0	0	0	0.65	0.32	0.43	0.46	0.37	0.41	0.44	**0.43**	0.43	0.89	0.1	0.18
Other	0.64	1	0.78	0.74	0.87	**0.8**	**0.75**	0.84	0.79	0.71	0.64	0.68	0.71	0.88	0.78
Average	0.55	0.64	0.51	0.69	**0.71**	**0.69**	0.67	0.69	0.67	0.6	0.58	0.59	0.69	0.69	0.65

Table 10: Three label classification of agendas for action as agendas for cooperative, restrictive treatment or other

	TF-IDF for NB			TF-IDF for SVM			TF-IDF for *k*NN			TF-IDF for DT			TF-IDF for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
de-escalation	0	0	0	**0.72**	0.67	**0.69**	0.69	0.67	0.68	0.48	**0.8**	0.6	0	0	0
support	0.82	**1**	0.9	0.93	0.95	**0.94**	0.93	0.94	0.93	**0.95**	0.81	0.88	0.82	**1**	0.9
average	0.68	0.82	0.74	**0.89**	**0.9**	**0.89**	**0.89**	0.89	**0.89**	0.87	0.81	0.83	0.68	0.82	0.74
	TF-IDF & data origin for NB			TF-IDF & data origin for SVM			TF-IDF & data origin for *k*NN			TF-IDF & data origin for DT			TF-IDF & data origin for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
de-escalation	0	0	0	0.63	0.75	0.68	0.59	0.67	0.62	0.54	0.75	0.62	0.64	0.67	0.65
support	0.82	1	0.9	0.94	0.91	0.92	0.93	0.9	0.91	0.94	0.86	0.9	0.93	0.92	0.92
average	0.68	0.82	0.74	0.89	0.88	0.88	0.87	0.86	0.86	0.87	0.84	0.85	0.88	0.88	0.88

Table 11: Classification of agendas classified as cooperative treatment in the step above as agendas for de-escalation or for support

	TF-IDF for NB			TF-IDF for SVM			TF-IDF for *k*NN			TF-IDF for DT			TF-IDF for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
Escalation	0.86	0.18	0.3	0.84	0.48	0.62	0.7	0.64	0.67	0.74	0.61	0.67	0	0	0
Punishment + ignorance	0.62	0.98	0.76	0.71	0.93	**0.81**	0.75	0.8	0.77	0.75	0.84	0.79	0.58	**1**	0.73
Average	0.72	0.64	0.56	0.77	**0.74**	0.73	0.73	0.73	0.73	0.74	**0.74**	**0.74**	0.33	0.58	0.42
	TF-IDF & data origin for NB			TF-IDF & data origin for SVM			TF-IDF & data origin for *k*NN			TF-IDF & data origin for DT			TF-IDF & data origin for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
Escalation	**0.88**	0.21	0.34	0.76	0.48	0.59	0.71	0.61	0.66	0.63	0.58	0.6	0.59	**1**	0.74
Punishment + ignorance	0.63	0.98	0.77	0.7	0.89	0.78	0.74	0.82	0.78	0.71	0.76	0.73	**1**	0.49	0.66
Average	0.73	0.65	0.59	0.73	0.72	0.7	0.73	0.73	0.73	0.68	0.68	0.68	**0.83**	0.71	0.69

Table 12: Classification of agendas classified as restrictive in the step above as agendas for escalation or punishment + ignorance

	TF-IDF for NB			TF-IDF for SVM			TF-IDF for *k*NN			TF-IDF for DT			TF-IDF for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
Punishment	0.72	**1**	0.84	0.74	0.97	0.84	**0.78**	0.97	**0.86**	0.76	0.76	0.76	0.72	**1**	0.84
Ignorance	0	0	0	0.67	0.15	0.25	**0.8**	0.31	**0.44**	0.38	**0.38**	0.38	0	0	0
Average	0.51	0.72	0.6	0.72	0.74	0.67	**0.79**	**0.78**	**0.75**	0.65	0.65	0.65	0.51	0.72	0.6
	TF-IDF & data origin for NB			TF-IDF & data origin for SVM			TF-IDF & data origin for *k*NN			TF-IDF & data origin for DT			TF-IDF & data origin for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
Punishment	0.72	**1**	0.84	0.71	0.91	0.8	0.71	0.91	0.8	0.71	0.73	0.72	0.72	**1**	0.84
Ignorance	0	0	0	0.25	0.08	0.12	0.25	0.08	0.12	0.25	0.23	0.24	0	0	0
Average	0.51	0.72	0.6	0.58	0.67	0.61	0.58	0.67	0.61	0.58	0.59	0.58	0.51	0.72	0.6

Table 13: Classification of agendas classified as punishment + ignorance in the step above as such

	TF-IDF for NB			TF-IDF for SVM			TF-IDF for ***k*NN**			TF-IDF for DT			TF-IDF for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
Agendas for not doing	0	0	0	0.62	0.23	0.33	**0.83**	0.06	0.11	0.43	**0.55**	**0.48**	0.67	0.18	0.29
Other	0.86	1	0.92	0.88	0.98	**0.93**	0.86	**1**	**0.93**	**0.92**	0.88	0.9	0.88	0.98	**0.93**
Average	0.74	0.86	0.79	0.85	**0.87**	0.84	**0.86**	0.86	0.81	0.85	0.83	0.84	0.85	**0.87**	0.84
	TF-IDF & data origin for NB			TF-IDF & data origin for SVM			TF-IDF & data origin for *k*NN			TF-IDF & data origin for DT			TF-IDF & data origin for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
Agendas for not doing	0	0	0	0.61	0.28	0.39	0.8	0.09	0.16	0.34	0.39	0.36	0	0	0
Other	0.86	1	0.92	0.89	0.97	**0.93**	0.87	**1**	**0.93**	0.9	0.87	0.88	0.86	**1**	0.92
Average	0.74	0.86	0.79	0.85	**0.87**	**0.85**	**0.86**	**0.87**	0.82	0.82	0.8	0.81	0.74	0.86	0.79

Table 14: Classification of agendas classified as other in the step above as agendas for not doing or other

	TF-IDF for NB			TF-IDF for SVM			TF-IDF for *k*NN			TF-IDF for DT			TF-IDF for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
Other	0.82	0.52	0.63	**0.83**	0.54	0.65	0.56	0.91	**0.69**	0.65	0.7	0.67	0.8	0.58	0.67
Multiclass	0.62	**0.88**	0.73	0.64	**0.88**	**0.74**	0.69	0.23	0.34	0.64	0.59	0.61	0.64	0.84	0.73
Average	0.73	0.69	0.68	**0.74**	0.7	0.69	0.62	0.58	0.52	0.64	0.64	0.64	0.72	0.7	0.7
	TF-IDF & data origin for NB			TF-IDF & data origin for SVM			TF-IDF & data origin for *k*NN			TF-IDF & data origin for DT			TF-IDF & data origin for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
Other	0.81	0.55	0.66	0.8	0.61	**0.69**	0.57	0.89	0.7	0.62	0.74	0.68	0.52	**1**	**0.69**
Multiclass	0.64	0.86	0.73	0.66	0.83	**0.74**	**0.7**	0.27	0.39	0.64	0.51	0.57	0	0	0
Average	0.73	0.7	0.69	0.73	**0.72**	**0.71**	0.63	0.59	0.55	0.63	0.63	0.62	0.27	0.52	0.36

Table 15: Classification of agendas classified as other in the step above as multiclass agendas or other

	TF-IDF for NB			TF-IDF for SVM			TF-IDF for *k*NN			TF-IDF for DT			TF-IDF for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
General	0	0	0	0.33	0.1	0.15	0.15	**0.67**	0.24	0.11	0.43	0.18	0.22	0.29	0.25
Other	0.93	**1**	**0.96**	0.93	0.98	**0.96**	**0.96**	0.69	0.81	0.94	0.73	0.82	0.94	0.92	0.93
Average	0.86	**0.93**	0.89	0.89	0.92	0.9	0.9	0.69	0.76	0.88	0.71	0.78	0.89	0.87	0.88
	TF-IDF & data origin for NB			TF-IDF & data origin for SVM			TF-IDF & data origin for *k*NN			TF-IDF & data origin for DT			TF-IDF & data origin for MLP		
	precision	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**	**precision**	**recall**	**f1**
General	0	0	0	**0.57**	0.19	**0.29**	0.12	0.43	0.19	0.15	0.43	0.22	0.5	0.1	0.16
Other	0.93	**1**	**0.96**	0.94	0.99	**0.96**	0.94	0.76	0.84	0.95	0.81	0.87	0.93	0.99	**0.96**
Average	0.86	**0.93**	0.89	**0.91**	**0.93**	**0.91**	0.88	0.74	0.8	0.89	0.78	0.82	0.9	**0.93**	0.9

Table 16: Classification of agendas classified as other in the step above as general agendas or other

Appendix 4

The overview of agenda for action types in *The NYT* and *The Guardian* in the period January until November 2013

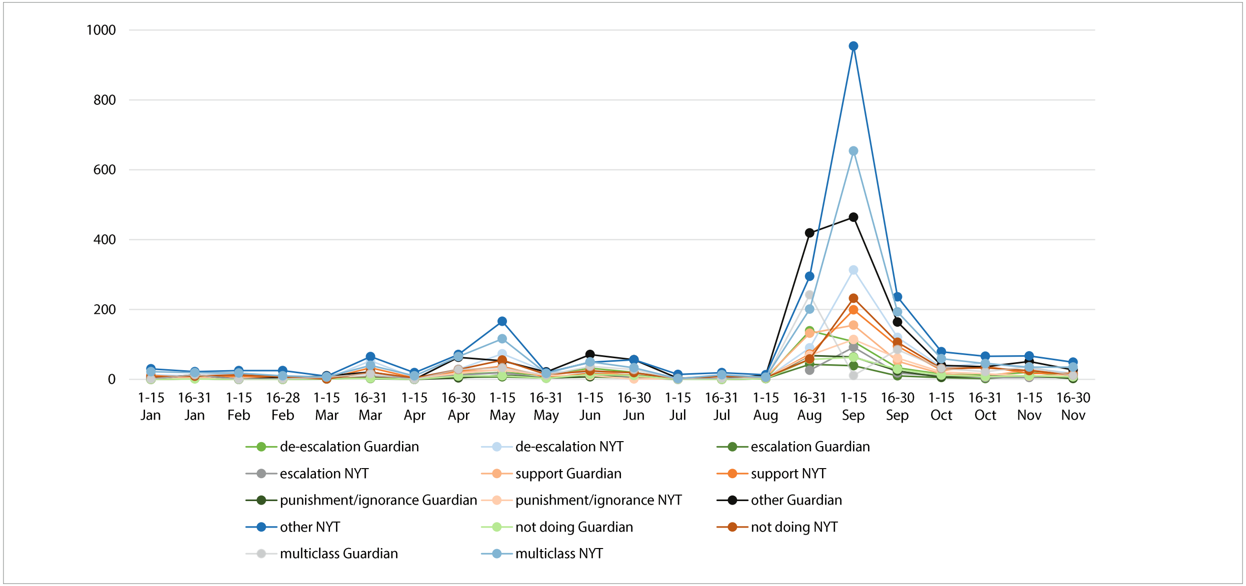


Figure 14: Agendas for action in *The NYT* and *The Guardian* from January until November 2013

Words can make people act. Indeed, a simple phrase *'Will you, please, open the window?'* can cause a person to do so. However, does this still hold, if the request is communicated indirectly via mass media and addresses a large group of people? Different disciplines have approached this problem from different angles, showing that there is indeed a connection between what is being called for in media and what people do. This dissertation, being an interdisciplinary work, bridges different perspectives on the problem and explains how collective mobilisation happens, using the novel term 'agenda for action'. It also shows how agendas for action can be extracted from text in an automated fashion using computational linguistics and machine learning. To demonstrate the potential of agenda for action, the analysis of *The NYT* and *The Guardian* coverage of chemical weapons crises in Syria in 2013 is performed.

Katsiaryna Stalpouskaya has always been interested in applied and computational linguistics. Pursuing this interest, she joined FP7 EU-INFOCORE project in 2014, where she was responsible for automated content analysis. Katsiaryna's work on the project resulted in a PhD thesis, which she successfully defended at Ludwig-Maximilians-Universität München in 2019. Currently, she is working as a product owner in the field of text and data analysis.

17,10 €
ISBN 978-3-95925-146-4