GSDRC Applied Knowledge Services
Governance · Social Development · Humanitarian · Conflict

www.gsdrc.org
helpdesk@gsdrc.org

Helpdesk Research Report

Methods for monitoring and mapping online hate speech

Brian Lucas

14.07.2014

Question

What models and methodologies exist to support online monitoring and mapping of hate speech and narratives of violence? How has monitoring hate speech been used to support programmatic activities?

Contents

1. Overview

Approaches to mapping hate speech[1] online can be classified into three principal groups based on their purpose:

- **Real time monitoring and mapping**: These projects, the best known of which is the Umati project in Kenya, aim to provide continuous monitoring of online media. Such projects are

[1] For legal purposes, hate speech is defined in national legislation. For research purposes, definitions can be varied and contested, but generally hate speech "refers to words of incitement and hatred against individuals based upon their identification with a certain social or demographic group. It may include, but is not limited to, speech that advocates, threatens, or encourages violent acts against a particular group, or expressions that foster a climate of prejudice and intolerance". (Gagliardone, Patel, and Pohjonen 2014, p. 5)

rare, but they have the potential to serve as early warning systems or enable a reaction to incidents as they occur.

- **Retrospective monitoring and mapping**: It has been more common to carry out analysis of online hate speech after it has happened by looking at archives of messages or collecting messages for a short time and then analysing them. Some of these projects have been pilot studies to test techniques for potential suitability for larger-scale use.
- **Discourse and content analysis**: These approaches examine potential hate messages within their social and political context to understand the meanings, motivations, and ideologies behind the messages, and to unpick the components of a message and its delivery. They do not aim to track trends in frequency or location, but to understand how hate messages are constructed and how they influence recipients. They are often labour-intensive, and are typically used on relatively small sets of data (comprising perhaps a few hundred messages) rather than for large-scale monitoring. (Gagliardone, Patel and Pohjonen 2014, pp. 19-22; Prentice et al. 2011)

Until recently, approaches to monitoring hate speech have relied on human analysts reading and classifying suspected messages, but attempts to apply automated techniques drawn from the field of corpus linguistics[2] are increasing. These approaches use large databases of texts, statistical methods, and machine learning to identify patterns and trends in language use. They have potential to process the massive amounts of data that can be collected through monitoring social media, and to operate in real time. However, they have so far had only limited success in dealing with the highly context-dependent nature of online hate speech. Linguistic features such as non-standard spelling and grammar, veiled or coded language, allusions, metaphors, slang, and the use of multiple languages make the challenge of accurately interpreting informal online speech difficult for computers, and even for humans. One project (Bartlett et al. 2014, p. 25) noted that even human analysts had to create a category for incomprehensible tweets, and most projects note that analysts do not agree on classifications for every suspect message.

Very few hate speech monitoring projects have been linked with programmatic activities to combat hate speech. During the 2013 Kenyan elections, the Umati project was linked with the Uchaguzi project which had a broader election monitoring mission and which referred instances of hate speech onwards to appropriate authorities. Most projects that we identified for this report only aimed to publicise and expose hate speech, or undertook after-the-fact analyses, and were not designed to respond to incidents.

[2] Corpus linguistics is an approach to studying language that is based on the analysis and comparison of large sets of language data called corpora (singular: corpus). A corpus is a collection of language (for example, a set of texts) that is representative of the way language is used in a particular context or community. (McEnery 2013)

2. Examples of real-time monitoring and mapping projects

Umati (Kenya)

Project website: http://www.ihub.co.ke/umati

Umati, a project on the Ushahidi[3] platform, monitored online hate speech in 2012 and 2013 in the run-up to Kenya's general elections in March 2013. It monitored selected blogs, forums, online newspapers, Facebook, and Twitter daily, in English and seven other languages. (iHub Research, 2013)

Umati relied on a manual process for collecting and categorising online hate speech. Six project workers scanned online platforms daily for hate and dangerous speech and recorded incidences in an online database. Messages were classified according to predefined characteristics depending on the influence of the author and their potential to incite violence, drawing on Benesch's (2013) framework for identifying dangerous speech. Incidences of particular concern were forwarded to Uchaguzi (see below) for action. (iHub Research, 2013)

Manual monitoring was important for assessing highly contextualised information in multiple languages. However, human error, especially due to fatigue, was a problem and scaling up the monitoring operation was expensive. (iHub Research, 2013, pp. 32-33) In future operations, the team intends to use Ushahidi's SwiftRiver software platform to assist with automatically monitoring and tagging messages. (iHub Research, 2013, p. 33)

Uchaguzi (Kenya)

Project information: http://blog.ushahidi.com/2013/02/11/uchaguzi-kenya-2013-launched/

Uchaguzi-Kenya was a project on the Ushahidi platform that enabled citizens to report problems occurring during Kenya's 2010 constitutional referendum and 2013 general election. It aimed to act as an early warning system and prevent the escalation of incidents. Other deployments have also taken place in Tanzania, Uganda, and Zambia in 2010 and 2011. (Omenya, 2013, pp. 9-10)

Uchaguzi included dangerous speech, rumours, and mobilisation toward violence among the threats it monitored, alongside other issues related to security, polling station management, and vote counting and reporting. (Chan, 2012; Ushahidi community, 2013) Kenyans could send reports via SMS, Twitter, Facebook, email, or via the Uchaguzi website. (Omenya, 2013, p. 19) The project staff was divided into teams which received and recorded reports from the public and from project colleagues, plotted reports on maps, translated messages, verified incoming reports with workers on in the field, relayed urgent messages to appropriate agencies for action, and carried out overall analysis and reporting. (Omenya, 2013, p. 25)

Uchaguzi has been considered largely successful in project evaluations (Chan, 2012; Omenya, 2013), but some areas for improvement have been suggested. The project had links with civil society organisations and government bodies, but many of these links were not well-organised and communications were irregular (Omenya, 2013, p. 15). This meant that although reports about threats

[3] Ushahidi began a project to map reports of election-related violence in Kenya in 2008, which has since expanded to become a non-profit organisation developing and deploying technology platforms for citizen participation in humanitarian and governance projects worldwide. (Source: www.ushahidi.com)

of violence were forwarded to appropriate agencies, feedback loops were not in place to confirm what actions were taken in response to reports. (Chan, 2012, pp. 14-16) The 2013 deployment generally suffered from late development and launch, and some technical problems hampered effectiveness. (Omenya, 2013, p. 20) Project volunteers were generally effective and efficient, but there were some problems in organising workflows efficiently. (Omenya, 2013, pp. 23-27)

Media Monitoring Project Zimbabwe

Project website: http://www.mmpz.org/

The Media Monitoring Project Zimbabwe is an independent trust launched in 1999 to promote freedom of expression and responsible journalism in Zimbabwe. It publishes monthly reports citing instances of hate speech in print media, electronic mass media, and social media, as well as thematic reports around elections, youth, corruption, and other issues. The most recent report on hate speech available from their website is dated January 2014. (Gagliardone et al., 2014, p. 21; Media Monitoring Project Zimbabwe, 2014)

3. Examples of retrospective monitoring and mapping projects

DEMOS study of anti-social media (Twitter, global)

Project report: http://www.demos.co.uk/publications/antisocialmedia

The think-tank Demos published a study in 2014 that examined the prevalence and patterns of use of racial and ethnic slurs on Twitter and tested the potential of automated monitoring of online speech. The study team collected publicly available tweets that contained one or more ethnic slurs based on a list of offensive terms compiled by the Wikipedia community. The study ran for nine days in 2012 and examined 126,975 tweets. (Bartlett et al. 2014, pp. 5-6)

A machine-learning programme called the Agile Analysis Framework was used to examine the potential for automated classification of tweets. Researchers developed a categorisation scheme and manually classified a sample of the tweets to create an initial training set which the computer analysed for correlations with linguistic features in the texts. The computer classified the remaining tweets, with researchers reviewing and re-training the computer's classification choices. Tweets were classified in four stages that assessed how suspected ethnic slurs were used in context, including differentiating between personal attacks and ideological statements. The computer was found to be fairly reliable in identifying ethnic groups targeted in messages that targeted ethnic groups, correctly classifying messages 75 per cent to 79 per cent of the time. However, performance in classifying messages as inflammatory or not was poor: only 54 per cent of the messages classified by the researcher as inflammatory were also identified as such by the computer, and only 57 per cent of the messages identified as inflammatory by the computer were also considered inflammatory by the researcher. (Bartlett et al. 2014, pp. 14-21)

The study also undertook a manual study of different types of usage of ethnic slurs, ranging from expressing negative stereotypes to explicit calls to action. The study team found that different analysts often disagreed on the interpretation of individual tweets, due to the wide range of types of usage,

multiple usages within a single tweet, ambiguousness of terms, and the cultural backgrounds of the analysts. (Bartlett et al. 2014, pp. 23-29)

Geography of Hate, Humboldt State University (USA)

Project website: http://users.humboldt.edu/mstephens/hate/hate_map.html

The Geography of Hate map is a demonstration project by Humboldt State University which shows the geographic distribution of tweets originating in the United States in 2012 and 2013 containing hate speech. The map was created by extracting tweets which contained specified "hate words" from the DOLLY Project (Digital OnLine Life and You) database at the University of Kentucky (see discussion of the DOLLY project below) and then having researchers read and classify each tweet individually as positive or negative in sentiment. The number of hateful tweets was aggregated at the county level and normalised by the amount of Twitter traffic. (Stephens, 2013a, 2013b)

Network of Social Mediators (Kyrgyzstan)

Project report: http://www.media-diversity.org/en/additional-files/documents/Hate-Speech-in-the-Media-and-Internet-in-Kyrgyzstan_English.pdf

The Network of Social Mediators, a Kyrgyz NGO, analysed content of state-run and private newspapers and online media, and selected Facebook and Twitter accounts, during two periods in 2013. Sources were monitored in the Kyrgyz, Russian and Uzbek languages. The analysis examined the role of local media in instigating or mitigating conflict following incidences of ethnic violence that took place in 2010 between Kyrgyz and Uzbeks. (Sikorskaya, 2014)

During the periods of analysis, sources were monitored five times per week and texts selected for analysis based on the presence of predefined keywords. Selected texts were classified by genre (news, analysis, opinions, interviews), by tone (propaganda, critical, neutral, positive, scientific), references to ethnicity, types of accusations made against the targets of hate speech, and other characteristics of the content of the texts. The project report does not contain details of the technologies or techniques used. (Sikorskaya, 2014)

Mouvement contre le racisme et pour l'amitié entre les peuples (France)

Project website: http://www.mrap.fr/

The *Mouvement contre le racisme et pour l'amitié entre les peuples* (MRAP, English: Movement Against Racism and for Friendship among Peoples) traces its roots back to organisations resisting anti-Semitism in Second World War France, and has since extended its work to supporting human rights and anti-racism efforts worldwide. (MRAP 2014)

An extensive study by MRAP in 2009 catalogued approximately 500 French-language web sites and more than 2,000 specific URLs promoting racist ideologies. These included organised hate groups' web sites as well as forums, blogs, and social networking sites. The researchers examined websites individually, identified recurrent themes and patterns of speech that were characteristic of different movements, and catalogued links to and from the studied websites to identify networks of hate sites. Some web sites were found to be overtly racist, while others used more subtle allusions or "coded" language. Openly racist organisations' websites (which would contravene French law) were often

hosted in countries without anti-racist legislation. (MRAP 2009; British Institute of Human Rights 2012, p. 29)

Institute of Human Rights and the Prevention of Xenophobia (Ukraine)

Project website: http://www.ihrpex.org/en

The Institute of Human Rights and the Prevention of Xenophobia (IHRPEX) is a "non-profit scientific and educational organisation" promoting human rights in Ukraine. (IHRPEX n.d.) The Institute carried out a study in 2011 that examined online aggressive, offensive, and threatening speech in 20 of the most popular Ukrainian social and political news websites. The study included content analysis of published articles and website users' comments, and a survey of website users about attitudes towards hate speech. The researchers noted that interpreting and coding comments was difficult and required extensive understanding of all of an author's comments and of the context of a discussion. Researchers collected a random sample of articles from the studied websites each day for five days. About one in three comments were considered hateful, although this study included comments directed at politicians and at other participants in online discussions individually as hateful. Hateful comments directed at ethnic groups made up 10 per cent of comments, and hateful comments against people from specific regions of Ukraine made up 20 per cent of comments. (IHRPEX 2011)

4. Discourse and content analysis techniques

Corpus linguistic approaches

Until recently, research into extremist ideology has tended to be qualitative in nature, relying on the reading and analysis of texts by researchers, and therefore has been limited in scale. There is, however, increasing interest in corpus linguistic techniques for automated processing of messages to help understand concepts and ideologies expressed in hateful texts. (Prentice et al. 2012, p. 259)

A study by Prentice et al. (2012) of 250 texts advocating violence on behalf of various Islamic extremist causes, written between 1996 and 2009, used the WMatrix[4] corpus analysis and comparison tool to analyse parts of speech and semantic elements. The programme identified words and concepts that emerge more frequently in the studied texts than in a corpus of general English language, presented as "word clouds" and "key concept clouds". The programme also made it possible to identify instances where extremist language appeared in conjunction with names of people and places, which may potentially help in mapping extremist networks. (Prentice et al. 2012, p. 281)

A similar example is a research project undertaken by Andrew Brindle (2009) which used the computer programme WordSmith[5] to analyse messages posted on a white supremacist web forum (Stormfront) in the USA. The programme identifies words and phrases that appear unusually often in comparison with other texts, and analyses how words and phrases of interest appear together in the suspect texts. The study combined this corpus-linguistic analysis with critical discourse analysis of a small sample of the messages. Both approaches helped develop an understanding of extremist authors' ideologies,

[4] http://ucrel.lancs.ac.uk/wmatrix/
[5] http://www.lexically.net/wordsmith/

the strategies they use to represent and argue for their views, the specific issues that were most important to them, and the range of positions taken by different members of the forum.

A study by Warner and Hirschberg (2012) from Columbia University examined text taken from Yahoo! News group posts and from a set of 452 suspected anti-Semitic websites to test an approach to automated classification of text. The study authors suggest that hate speech employs well-known stereotypes to convey its messages, and that each stereotype (or target of hate) "has a language all its own, with one-word epithets, phrases, concepts, metaphors and juxtapositions that convey hateful intent" (p. 21). The authors generated a classifier for anti-Semitic speech drawing on known stereotypes. The task was further complicated by the fact that some messages posted on discussion forums used deliberate misspellings and other techniques to evade simple filters. In the study, human classifiers were able to correctly identify hateful messages as hateful 59 per cent of the time, and correctly identifying non-hateful messages as benign 68 per cent of the time. The computer achieved a level of success similar to the human analysts, at 68 per cent and 60 per cent.

Content and composition analysis

In an analysis of online extremist texts written between 2000 and 2009 that incited violence in Gaza and the West Bank, Prentice et al. (2011) combine content analysis and semantic analysis in an overall approach they call Content and Composition Analysis. The project was a rigorous academic study aiming to identify techniques that extremist authors used to influence readers and to observe how these approaches changed after Israeli military action in 2008-09.

The content analysis component involved researchers reading texts to identify occurrences of "persuasive devices" that extremist authors used to influence readers. In this study, nine persuasive devices were examined: direct pressure, exchanging, persuasion, upward appeals, social proof, moral proof, activation of commitments, inspirational appeals, and liking (definitions are provided in Appendix 1). Authors of extremist texts examined in this study relied most heavily on moral proof (appeals to justice and morality), social proof (appeals to social and cultural comparisons and values), and upward appeals (citing authority figures). The semantic analysis component of this analysis used the computer programme WMatrix to identify concepts that appeared in the studied texts significantly more often than in 'normal' texts. It also identified how concepts occurred together in texts, and revealed trends in the appearance of these concepts over time.

Combining manual and automated methods in this study enabled researchers to identify distinctive features of hate messages, identify different strategies that authors with different affiliations used to influence readers, and track how the content of messages changed over time in response to events.

5. Datasets useful for supporting hate speech monitoring

DOLLY

Project website: http://www.floatingsheep.org/

The DOLLY project (Digital OnLine Life and You) based at the University of Kentucky in the USA continuously monitors Twitter in real time. It stores eight million tweets per day (more than three billion tweets in total) and does some basic analysis, indexing, and geocoding. A few research projects

have used the database for various purposes (see for example Geography of Hate, above). The project intends to develop a user-friendly front-end to enable easier access for researchers.

Hatebase

Project website: http://www.hatebase.org

Hatebase is an online repository of samples of hate speech in multiple languages, intended to assist organisations and researchers in predicting violence. It is an initiative under the Sentinel Project for Genocide Prevention. It offers two main features: a Wikipedia-like interface which allows users to classify and record location-specific hate speech, and an Application Programming Interface (API) that allows developers to connect Hatebase data with other tools to predict conflict and genocide. Data collected through Hatebase may be used in conjunction with other warning factors to provide insights into when speech may turn into action.

6. Websites that collect reports from the public

There are a variety of websites that allow members of the public to file reports of hate speech incidents. Such initiatives are unlikely to be comprehensive or systematic, and cannot give an accurate picture of the extent of incidents (British Institute of Human Rights 2012, p. 29) so we have not attempted to catalogue them in detail for the purposes of this report. It is also often unclear what analysis or action is taken as a result of such reports. Two examples include:

INHOPE – International Association of Internet Hotlines (http://www.inhope.org): Acts as a gateway to enable members of the public to report hate speech and other potentially illegal online content to the appropriate national authorities. The project is operated by the European Commission and international law enforcement agencies.

Hate Speech Watch (http://www.nohatespeechmovement.org/hate-speech-watch/): Website visitors can file reports about hate speech incidents and project staff prepare monthly messages which aim to respond to interests expressed by the online community. The project is operated by youth volunteers and is supported by the Council of Europe.

7. About this report

Contributors

We would like to thank the following experts for suggesting projects and literature for inclusion in this report:

- Tony McEnery, Professor, Department of Linguistics and English Language, Lancaster University, UK
- Paul J. Taylor, Professor, Department of Psychology, Lancaster University, UK
- Andrew Brindle, Assistant Professor, Department of Applied English, St. John's University, Taiwan R.O.C.
- Paul Iganski, Professor of Criminology & Criminal Justice, Lancaster University
- Caroline Sugg, Head of Special Projects, Advisory & Policy Team BBC Media Action, UK

Bibliography

Bartlett, J., Reffin, J., Rumball, N., and Williamson, S. (2014) *Anti-Social Media*. Demos. http://www.demos.co.uk/publications/antisocialmedia

Benesch, S. (2012). *Dangerous Speech: A Proposal to Prevent Group Violence*. The Dangerous Speech Project. http://www.worldpolicy.org/sites/default/files/Dangerous%20Speech%20Guidelines%20Benesch%20January%202012.pdf

Brindle, A. (2009) *A Linguistic Analysis of a White Supremacist Web Forum*. PhD dissertation (unpublished). Lancaster University.

British Institute of Human Rights (2012) *Mapping study on projects against hate speech online*. Council of Europe. http://www.coe.int/t/dg4/youth/Source/Training/Training_courses/2012_Mapping_projects_against_Hate_Speech.pdf

Chan, J. (2012). *Uchaguzi: A Case Study*. Harvard Humanitarian Initiative and Knight Foundation. http://www.knightfoundation.org/media/uploads/media_pdfs/uchaguzi-121024131001-phpapp02.pdf

Gagliardone, I., Patel, A., and Pohjonen, M. (2014) *Mapping and Analysing Hate Speech Online: Opportunities and Challenges for Ethiopia*. University of Oxford and Addis Ababa University. http://pcmlp.socleg.ox.ac.uk/sites/pcmlp.socleg.ox.ac.uk/files/Ethiopia%20hate%20speech.pdf

IHRPEX (Institute of Human Rights and the Prevention of Xenophobia) (n.d.) *Concept*. IHRPEX website. http://www.ihrpex.org/en/page/1/concept

IHRPEX (Institute of Human Rights and the Prevention of Xenophobia) (2011) *Phenomenon of the cyber-hatred in the Ukrainian Internet space: Summary of the report*. Institute of Human Rights and the Prevention of Xenophobia. http://www.ihrpex.org/en/article/2086

iHub Research. (2013). *Umati Final Report*. Ushahidi and iHub Research. http://www.research.ihub.co.ke/uploads/2013/june/1372415606__936.pdf

McEnery, T. (2013) Corpus: Some Key Terms. CASS Briefing no. 1. ESRC Centre for Corpus Approaches to Social Science (CASS), Lancaster University, UK. http://cass.lancs.ac.uk/?page_id=956

MMPZ (Media Monitoring Project Zimbabwe) (2014a) Media Monitoring Project Zimbabwe. http://www.mmpz.org/

MMPZ (Media Monitoring Project Zimbabwe) (2014b) Hate Speech Report for the month ending January 2014. Media Monitoring Project Zimbabwe. http://www.mmpz.org/hate-language/hate-speech-january-2014

MRAP (Mouvement contre le racisme et pour l'amitié entre les peuples) (2009) *Internet, enjeu de la lutte contre le racisme*. Mouvement contre le racisme et pour l'amitié entre les peuples. http://www.mrap.fr/documents-1/rapport-mrap2009.pdf

MRAP (Mouvement contre le racisme et pour l'amitié entre les peuples) (2014) *Meet the MRAP*. MRAP website. Mouvement contre le racisme et pour l'amitié entre les peuples. http://www.mrap.fr/english/meet-the-mrap

Omenya, R. (2013) *Uchaguzi Kenya 2013: Monitoring & Evaluation*. iHub Research and HIVOS. http://www.ihub.co.ke/ihubresearch/jb_UchaguziMEFinalReportpdf2013-7-5-14-24-09.pdf

Prentice, S., Taylor, P. J., Rayson, P., Hoskins, A., O'Loughlin, B. (2011). "Analyzing the semantic content and persuasive composition of extremist media: A case study of texts produced during the Gaza conflict", *Information Systems Frontiers*, 13, 61-73. http://dx.doi.org/10.1007/s10796-010-9272-y

Prentice, S., Rayson, P., and Taylor, P. J. (2012). "The language of Islamic extremism: Towards an automated identification of beliefs, motivations and justifications", *International Journal of Corpus Linguistics*, 17(2), 259-286. http://dx.doi.org/10.1075/ijcl.17.2.05pre

Sikorskaya, I., and Gafarova, S. (2014). *Hate Speech in the Media and Internet*. School of Peacemaking and Media Technology in Central Asia. http://www.media-diversity.org/en/additional-files/documents/Hate-Speech-in-the-Media-and-Internet-in-Kyrgyzstan_English.pdf

Stephens, M. (2013a). *The Geography of Hate*. http://www.floatingsheep.org/2013/05/hatemap.html

Stephens, M. (2013b) *Geography of Hate: Geotagged Hateful Tweets in the United States*. Humboldt State University. http://users.humboldt.edu/mstephens/hate/hate_map.html

Uchaguzi (2013) *Uchaguzi Kenya 2013: Monitoring & Evaluation.* iHub Research. http://www.ihub.co.ke/ihubresearch/jb_UchaguziMEFinalReportpdf2013-7-5-14-24-09.pdf

Umati (2013) *Umati Final Report*. iHub Research. http://www.research.ihub.co.ke/uploads/2013/june/1372415606__936.pdf

Ushahidi community (2013) *Ushahidi wiki: Uchaguzi - Kenyan Elections 2013*. https://wiki.ushahidi.com/display/WIKI/Uchaguzi+-+Kenyan+Elections+2013

Warner, W., and Hirschberg, J. 'Detecting Hate Speech on the World Wide Web', *Proceedings of the 2012 Workshop on Language in Social Media (LSM 2012)*, pp, 19-26, Montreal, Canada. Association for Computational Linguistics. http://dl.acm.org/citation.cfm?id=2390377

Suggested citation

Lucas, B. (2014) *Methods for monitoring and mapping online hate speech*. GSDRC Helpdesk Research Report no. 1121. University of Birmingham.

Appendix 1: Definitions and examples of nine persuasion behaviours

Table 1 Definitions and examples of nine persuasion behaviors as a function of theoretical categorization

Motivational Frame	Power Use	Imagery	Tactic	Definition	Example
Argument-related (Instrumental)	Hard	N/A	Direct Pressure	Pressure tactics that include commands, demands, forceful assertiveness, intimidation, and threats	"Raise your arms and fight to escape from this humiliation and shame!"
	Mid	N/A	Exchanging	Explicit or implicit promise that you will receive rewards or tangible benefits if you comply with a request or support a proposal. Can be associated with 'scarcity'	"When the enemy targets our women and children we should target theirs"
	Soft	N/A	Persuasion	Argument that attempts to explain reasons, or presents information in support of a position. Includes (but not limited to) the use of logical arguments, factual evidence, and statements of 'expertise' (i.e., Because that's the nature of things)	"Another obstacle is the need for advanced means of resistance to counter the occupation and defend the people and the land"
Audience-related (Relational)	Hard	Religion	Upward Appeals	Use of authority-based comparisons, which seeks to persuade you that a higher authority approves the action, or which link an issue idea or cause to another positive concept associated with an authority.	"Shaykh Ibn al Uthaymeen says: If the enemy kill our women and children it appears to me that we are allowed to kill their women and children"
	Mid	Society	Social Proof	Use of social comparisons, typically to groups, communities or societies rather than people (see above) in support of a viewpoint or argument. Includes comparison to cultural values, whereby the speaker indicates the value of art, history or traditions of one's own group	"And our people in Palestine totally reject such a description along with our Arab and Muslim peoples"
	Soft	Morality	Moral Proof	Use of moral comparisons, either justifying the morality of a particular position or action, or highlighting immorality in the actions or positions of an out-group, including suggestions of double standards.	"If a herd of dogs and pigs had suffered a tenth of what the Palestinians in Gaza have suffered, all institutions of the non-believer West would have risen in protest"
Speaker-related (Identity)	Hard/Mid	Moral-Social (group unity)	Activation of commitments	Messages that remind listeners of their commitment to a position, group or action, or suggestion of a debt owed because of past events or actions of others. This can include arguments around building a coalition or single voice.	"It is incumbent upon us to use all our resources to confront the attack on our ummah"
	Hard/Mid	Egoistic	Inspirational appeals	Use of an emotional request or proposal that arouses enthusiasm by appealing to positive or negative self-feeling (e.g., you will feel better about yourself if you comply), altruism (e.g., I need your compliance very badly), or esteem (e.g., people will think better of you if you comply).	"And we will be the coming power insha' Allah"
	Soft	Moral-social	Liking	Use of friendly or helpful messages by a speaker to put the listener in good frame of mind (e.g., Ingratiation). This might include recognizing the struggle of a particular group, or indicating allegiance with a group to improve credibility.	"For those who asked that I reconsider my view on this, I promise I will review it again"

Source: Prentice et al. 2011, p. 65