



Article

Press and social media reaction to ideologically inspired murder: The case of Lee Rigby

Discourse & Communication

2015, Vol. 9(2) 237–259

© The Author(s) 2015

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/1750481314568545

dcm.sagepub.com**Tony McEnery**

Lancaster University, UK

Mark McGlashan

Lancaster University, UK

Robbie Love

Lancaster University, UK

Abstract

This article analyses reaction to the ideologically inspired murder of a soldier, Lee Rigby, in central London by two converts to Islam, Michael Adebowale and Michael Adebola. The focus of the analysis is upon the contrast between how the event was reacted to by the UK National Press and on social media. To explore this contrast, we undertook a corpus-assisted discourse analysis to look at three periods during the event: the initial attack, the verdict of the subsequent trial and the sentencing of the murderers. To do this, we constructed and analysed corpora of press and Twitter coverage of the attack, the conviction of the suspects and the sentencing of them. The analysis shows that social media and the press are intertwined, with the press exerting a notable influence through social media, but social media not always being led by the press. When looking at social media reaction to such an event as this, analysts should always consider the role that the press are playing in forming that discourse.

Keywords

Corpus analysis, ideology, media, murder, twitter

Corresponding author:

Tony McEnery, Department of Linguistics and English Language, Lancaster University, Lancaster, LA1 4YW, UK.
Email: a.mcenery@lancaster.ac.uk

Introduction

Lee Rigby, a fusilier in the British army, was attacked by two men with knives and killed on the streets of Woolwich, London, shortly after 14:00 on 22 May 2013. The murder attracted a great deal of media attention because it was particularly brutal and made intentionally public – Rigby was first run down by a car and then subjected to an attempted beheading by two men: Michael Adebolajo and Michael Adebawale. The men chose to attack Rigby because he was wearing a sweatshirt displaying with a logo for the charity organisation Help for Heroes which raises money for war veterans in the United Kingdom. The decision to openly commit violent murder was a provocative one. It was intended to grab the attention of passers-by and onlookers, a number of whom recorded the incident on mobile smart phones and/or reported it using not only the emergency services but also social media and networking sites such as Twitter and Facebook. In video footage, Adebawale addressed onlookers filming the incident, delivering the following message:

The only reason we have killed this man today is because Muslims are dying daily by British soldiers and this British soldier is one; he is an eye for an eye and a tooth for a tooth.

The incident became a headline news story, and the development of the legal case, from the attack through to the eventual conviction and sentencing of Adebolajo and Adebawale for the murder of Lee Rigby, has been widely reported online and in the mainstream international press. Adebolajo was given a whole-life sentence, meaning he will never be released from prison. Adebawale must serve a minimum of 45 years in prison before being considered for release.

The event warrants analysis as it was ideologically inspired: Adebolajo and Adebawale wanted to use the murder to publicise their belief that British soldiers were killing Muslim civilians and that this must stop. In assessing the extent and ways in which discourses around this ideologically inspired murder were realised in society, we could simply look at press reports. However, we argue that a press-only analysis of such an event is limited due to the fact that many people now engage with news stories in a more interactive way via social media. Hence, an analysis of social media in this case seems necessary. An analysis of social media and press reaction is also of interest as while the latter is often studied using corpus-assisted discourse studies, the former is not (although see Baker and McEnergy's in press study of discourses on Twitter around the television programme Benefits Street). Consequently, this article will reflect upon how a corpus-based discourse analysis of social media (specifically Twitter) may be undertaken.

In this study, we take a corpus-assisted discourse studies approach (Partington et al., 2004). Methodologies combining the use of *corpus linguistics* (CL; 'a methodology rather than an independent branch of linguistics'; McEnergy et al., 2006: 7) and *critical* approaches to the study of language, including critical discourse analysis (CDA), have become 'increasingly popular in critical approaches to discourse analysis' and have proven fruitful to critical research (cf. Baker et al., 2008). In the tradition of CDA, practitioners drawing on a combined CDA/CL methodology are interested in the discourses that permeate the relationships among social groups, particularly those involving imbalanced power dynamics. CDA/CL studies have therefore sought to analyse discourse in various language datasets, often drawn from press reporting, with regard to topics such

as same-sex relationships (Bachmann, 2011), anti-terror laws (Mautner, 2009), refugees and asylum seekers (Baker and McEnery, 2005) and representations of minorities such as the Muslim community in the United Kingdom (Baker, 2010; Baker et al., 2013; Gabrielatos and Baker, 2008).

In using tools from CL, the article takes both a contrastive and diachronic approach to the event. This diachronic dimension will allow us to explore how, over time, the reporting of the event developed on both social media and in the press. It also allows us to see whether the contrast between the two remains static or whether there are changes over time. In the following sections, we will present a brief background to the event, outline the data we use in our study, explore the data linked to each of the three periods studied and, finally, draw conclusions from our analysis.

Corpus collection

The data are gathered from two separate sources, split into three periods. As noted, these are the *attack* on Lee Rigby, the *conviction* of Michael Adebolajo and Michael Adebowale of the murder of Lee Rigby and the final *sentencing* of Adebolajo and Adebowale in court. One data source was Twitter, from which we collected tweets which occurred around the time of these events, and the other was newspaper reportage of the case. Twitter data were gathered using DataSift, an online service enabling access to historic Twitter data, and newspaper articles were collected using LexisNexis, an online news aggregator service.

In order to make data comparable, we collected both the newspaper and Tweet corpora by using the following queries. We looked for any tweets/articles containing the following (case insensitive) terms in their entirety:

Adebowale OR Adebolajo OR Lee Rigby OR Woolwich

The search term *Woolwich* was used as we noticed that the terms had become used in a metonymic sense for the attack, rather than simply the location itself.¹ The other search terms are the names of the victim *Lee Rigby* and of the perpetrators of the attack *Adebowale* and *Adebolajo*. In considering the press, we deemed these search terms to be sufficient to capture most, if not all, of the articles published in the UK national press which discussed the attack. As is apparent in Table 1, the terms returned a substantial dataset.

Nexis News Search (LexisNexis, 2014) was used to collect all the articles from the British national press that make up the newspaper part of our dataset. Using Nexis, we searched only articles that appear in UK national newspapers, both printed and online. The following titles were collected: *Daily Mail*, *Daily Star*, *Daily Star Sunday*, Express Online, guardian.co.uk, i – Independent Print Ltd, *The Independent on Sunday*, independent.co.uk, *Mail on Sunday*, *MailOnline*, *Morning Star*, telegraph.co.uk, *The Business*, *The Daily Telegraph* (London), *The Express*, *The Guardian* (London), *The Independent* (London), *The Mirror*, *The Observer*, *The People*, *The Sunday Mirror*, *The Sunday Telegraph* (London), *The Sunday Times* (London) and *The Times* (London).

Data gathered from these sources were used to construct a set of corpora about Lee Rigby's murder. We have used an initial letter system to refer to different types of

Table 1. Total number of articles and words in each sub-corpus of press coverage about the Lee Rigby murder case.

Event and medium	Time period	Articles	Words
Attack Press (AP)	22–28 May 2013	1772	1,310,625
Conviction Press (CP)	19–25 December 2013	307	251,611
Sentencing Press (SP)	26 February–4 March 2014	156	146,879
Total for all press corpora		2235	1,709,115

corpora and distinguish between the period of the initial attack (A), the conviction (C) and the sentencing (S). We also make a distinction between data from the press (P) and Twitter (T). Finally, we note retweets as (R). So for the press corpora, each sample was taken in a period which begins with the event in question and runs for one week.

The use of Twitter material required us to be mindful of the features of the medium when undertaking a corpus analysis. Conversation on Twitter is built up from tweets, small texts limited to 140 characters of written language. Within tweets, users can *mention* (i.e. send messages to) others explicitly by including a username prefixed by the @ character (e.g. @CorpusSocialSci) and annotate their tweet using *hashtags* (any terms without spaces prefixed by the # character, e.g. #corpuslinguistics) that act, in essence, as metadata linking tweets containing the same hashtag. The hashtag itself is usually easily interpreted by a human user as reflecting a particular topic or theme, for example, #woolwich. Hashtags may also reflect other meanings, for example, sentiment, as in #ripleerigby. Once a tweet has been sent (*tweeted*), other users can then *retweet* and/or *favourite* and/or *reply* to that tweet. A *retweet* is a way for a user to forward or ‘re-post’ existing tweets to their followers and can extend the reach of a tweet far beyond the network of the original tweeter. A *favourite* is often used to ‘bookmark’ a tweet for later reference, but primarily people use favourites to notify the original poster of the tweet that they like the tweet. A *reply* is when a user responds to another user’s tweet using the reply button. Replies are like mentions in that the username of the user being mentioned can occur anywhere in the tweet prefixed by the @ character. In this article, usernames and retweets were prominent in the analyses. An initial investigation of an early version of the Tweet corpora revealed that hashtags were, in essence, reliable metadata tags allowing for the recovery of relevant tweets relating to the incident. So as well as the search terms used to build the press corpora, two hashtags, #woolwich and #leerigby, were included in the search term used to construct our Tweet corpora. As with the press corpora, the Tweet corpora for each event were collected from the week immediately following the incident in question. Unlike the press corpora, however, we divided the Tweet corpora for each event into two. The split is between original tweets and retweets. We undertook this split principally to facilitate consideration of the effects of retweeting, as discussed later. Table 2 shows the composition of the Tweet corpora.

One challenge we faced with the Twitter data was a high incidence of variant spelling forms, which has the effect of skewing frequency data. As variant spelling on social media was not the focus of our research study, we used a piece of software, VariAnt,² in order to normalise spelling in the corpora for the purpose of corpus analysis.

Table 2. The Lee Rigby tweet corpora.

	Attack Tweets (ATs)	Conviction Tweets (CTs)	Sentencing Tweets (STs)
Tweets	419,775	31,573	26,991
Words	6,851,577	553,543	459,879
	Attack Retweets (ARs)	Conviction Retweets (CRs)	Sentencing Retweets (SRs)
Retweets	596,217	30,817	31,027
Words	13,183,958	682,442	669,348

Analytical approach

Our analyses used three key methods from CL: collocation, word-sketches and keywords, undertaken using the online SketchEngine³ concordancing program.

Collocates are the words that occur ‘frequently within the neighbourhood of another word, normally more often than we would expect the two words to appear together because of chance’ (Baker et al., 2013: 36). In our analysis, collocates were calculated within a span of five words to the left and five words to the right of the search terms (see Baker et al., 2013: 36). The logDice measure (Rychlý, 2008) was used to assess the strength of collocation, as it has been shown to work well with corpora of different sizes, such as those used in this study (Rychlý, 2008: 9).

Collocation analysis is useful for our purposes as it enables us to *look around* the most frequent lexical items that exist in the corpora. They also help reveal discourse prosodies (Stubbs, 2001: 65), patterns of words with an inherently negative or positive evaluation being systematically associated with a word.

Word-sketches represent a helpful extension of collocation, linking collocates to frequent grammatical roles that they play in a corpus. So, for example, if a word collocates with another and appears strongly in the grammatical role of direct modifier to that word, then the word-sketch shows this. Similarly, a word-sketch can show whether a noun collocating with a verb has the propensity to appear as subject or object.

We used keywords to characterise the language of the articles and tweets in the three periods of reporting about the Lee Rigby murder and to provide a general overview of the data. Keywords are words which are the most unusually frequent in our data compared to another corpus (a reference corpus). In this case, we compared the press data with the Twitter data and vice versa. Calculating keywords in this way allowed us to see what it is about the language of the press reporting that characterises it against the Twitter data, and vice versa. In doing so, we could find out which concepts/topics were salient to the articles published and tweets retrieved about each of the murder, conviction and sentencing periods. In this article, the press corpora are also contrasted with a large corpus of general English, the 100-million-word British National Corpus (BNC). The tweets are also compared to a corpus of approximately 1,000,000 general tweets gathered in 2014. Each retweet corpus is contrasted with its corresponding tweet corpus – so the Attack Tweet (AT) is used as the reference corpus for the Attack Retweet (AR), for reasons discussed in the next section.

We use keywords in two ways. First, we were interested both in which elements of the discourse, realised through the keywords, remain relatively stable/consistent over time and in which elements of the discourse are unstable or specific to one event only. We use the keywords for each period as ‘lexical signposts’ (Baker, 2004: 90), pointing towards interesting words to consider for further analysis. This is achieved using a combination of manual concordance analysis and the calculation of collocates of some of the more frequent keywords. On the basis of such an analysis, we move to our second use of keywords – we organise keywords according to principal themes within the discourse. Hence, for example, in the analysis of the Attack Press (AP) corpus, the keywords *adebolajo*, *adebowale* and *muslims* are placed in the general discourse theme of ‘Identity of the perpetrators’ as these keywords, when collocates and concordances are consulted, predominantly place themselves in this role in the discourse. While a word typically appears in one discourse theme in our analysis, certain words link to more than one theme. In such cases, those words are placed in multiple themes and the reason for this is discussed in the analysis of that word.

In looking at keywords, we focused our analysis by examining the 20 most salient keywords in the press and Tweet corpora in the period. Our examination of retweets had a different focus – as is discussed in the next section.

Retweeting, social structure and ideology

The approach to retweets taken in this article is to explore what voices and themes become dominant in the discourse because of retweets. Hence, as noted, in this article we undertake a keyword analysis by contrasting the Tweet corpus of each incident with the Retweet corpus of each incident. By doing this, the repetitions in the retweet corpus reveal themselves easily as the frequency of various tweets which are present in the tweet corpus is amplified through retweeting in the retweet corpus. Approaching such an analysis, however, leads to the question ‘What should be analysed?’ The retweet corpus yields keywords which are lexical in nature but also much in the way of metadata – Twitter usernames, for example. On an exploration of the data, we decided that approaching the analysis through the metadata was most productive. This allows us to pick up main themes that are retweeted, whereas approaching those themes on a word-by-word basis is less efficient. For example, the following AT was retweeted 12,727 times:

@SoYouWannaPwn Half the things people are tweeting should put them in jail @metropolitanpoliceuk The news hasnt even got confirmed stories yet!

This is the most retweeted comment in the AR – yet if we approach the data by looking at words alone, this is not apparent because, although these words are repeated very often, other words are repeated in a range of tweets, meaning that this dominant tweet, while suggested by the top 10 lexical keywords in the AR corpus, is by no means clear in that list: *hasnt*, *half*, *friendliest*, *hospitable*, *stories*, *jail*, *tolerant*, *tweeting*, *confirmed*, *bosnia*.

Given that the goal of studying the retweet corpus is to look at what themes emerge as preferred from the discussion, implying an implicit evaluation of the tweeted discourse,

we chose instead to focus on the key users on Twitter who people chose to retweet. Consequently, only tweets receiving over 1000 retweets from the most retweeted users are included in the analyses in this article.

Extremely frequent retweets (i.e. those repeated hundreds or thousands of times) in particular have implications for the analysis and interpretation of frequency, hence our focus on them in this article. We would argue strongly that this is not a disadvantage of such data, however. Retweeting is particularly important both as a social practice and as a type of data for discourse analysis. When a tweet is retweeted, a particular user's linguistic choices (hence, ideological stance) are recast and repeated. Any eventual corpus of Twitter data that includes retweets will therefore include duplicate texts. A Twitter-based corpus which includes retweets is, we believe, useful for interpreting consensus of ideology and opinion, although if retweets are not taken into account, we may privilege the language of a few very frequent retweets at the expense of viewing more diverse and diffuse forms of language on the same topic. While Twitter allows anyone to broadcast their thoughts, there are resulting inequalities in bandwidth and distribution – the bandwidth of broadcast for a tweet is determined by the number of followers a user has and the number of retweets that a tweet achieves. Not all Twitter users are equal, and the social networks that following and retweeting represent are observable manifestations of social influence and agreement in ways analogous to newspaper distribution figures.

When people speak, write or tweet, they make – consciously or subconsciously – linguistic choices which are taken to be ideologically informed. Frequency is ‘one of the most central concepts underpinning the analysis of corpora’ (Baker, 2006: 47–48), which ‘can provide information about the sorts of concepts that are privileged in society’ (Baker, 2014: 75) – and frequency is precisely what is altered by retweeting. Choice of words expresses an ideological position (Stubbs, 1996: 107), and we argue that no retweet is neutral – amplification of frequency by retweeting allows us to take a frequency-based approach to identify which tweets are preferred by discourse communities and hence which ideological positions are retweeted. A simple retweet implies agreement typically, and consequently, the increased frequency of words from that tweet being retweeted is as a consequence of ideology. So retweets, we would argue, are not ‘noise’. They are direct evidence of the propagation of ideological views by members of a discourse community, and, as such, retweets should not be discarded – they are indicative of agreement, agreement signalled by repetition.

Analysis and discussion

In the following analyses, we look in turn at the discourse surrounding the attack, conviction and sentencing on both Twitter and in the UK national press. Throughout we compare and contrast the two, drawing out key similarities and differences between both. We look at how those discourses interact and develop across the three events covered by this study. For each time period, we also examine frequent retweets which demonstrate what elements of the Twitter discourse were selected by discourse communities on Twitter for propagation and amplification.

Table 3. The top 20 keywords in the AT and AP corpora organised by discourse theme.

Category	Tweet (AT) keywords	Press (AP) keywords
Condolence	<i>rip</i>	
Identity of the perpetrators	<i>british, muslims, suspect</i>	<i>adebolajo, adebowale, muslims</i>
Identity of the victim	<i>british, drummer, lee, rigby, soldier</i>	<i>drummer, lee, rigby, soldier</i>
Incident	<i>attack, incident, killed, killing, murder</i>	<i>attack, killing, murder, terror, terrorism</i>
Location of incident	<i>london, woolwich</i>	<i>woolwich</i>
Politics	<i>edl</i>	<i>cameron</i>
Problem	<i>british</i>	<i>choudary, islam</i>
Question	<i>happened</i>	
Reaction	<i>muslim</i>	<i>islam, muslim</i>
State reaction	<i>police</i>	<i>arrested, mi5, police</i>

edl: English Defence League; AT: Attack Tweets; AP: Attack Press.

It is interesting to note that the top hashtags in the AT also generally fall into these categories, being focussed upon location (#london, #woolwich), politics (#edl) and condolence (#rip). The exception is #racism, which will be discussed in the context of the AR (Attack Retweet corpus).

Reaction to the attack on Lee Rigby, 22–28 May 2013

Let us begin by considering the reaction to the attack on Lee Rigby on Twitter and in the press, summarised in Table 3 which shows the top 20 keywords for the AT corpus and the AP corpus. These keywords have been placed into groups reflecting the principal themes within the discourse.⁴

Many of the words in this table are self-explanatory, as are the discourse themes. People are tweeting and the press is writing about the identity of the victim, the incident, the location of the incident, the identity of the perpetrators and the state reaction to the event (through the police and intelligence services, principally Military Intelligence, Section 5 (MI5)).

More light on the differences between Twitter and press reaction comes from looking at where the categories differ across the table. The personal nature of Twitter is evident in that expressions of condolence are frequent enough on Twitter to enter the top 20 keywords in the Tweet corpus – this does not happen with the press data. The word *happened* has the same distribution as *rip* – it is salient in the AT but not the AP corpus. This is because the word is linked very strongly to a request for information – the strongest collocate of this word is *what*. In the word-sketch, this appears as a strong collocate as the subject of *happened*, in forms including *whats*, *wat* and *wtf*. In the context of the word's use, its absence from the news reportage is also understandable – the news is static and monologic relative to Twitter – people do not write to newspapers asking them to print stories about what happened. However, Twitter is dynamic and dialogic – people may respond rapidly to such requests for information.

In addition to differences, there are intersections in the analysis, notably between the identity of the victim, the identity of the perpetrators and a perceived problem as shown by the word *british*. In the analysis above, it is included as part of the identity of both and as a problem. An exploration of the collocates reveals why: within the top 10 collocates

of *british* in AT, there are three clear sets – those linking *british* to the victim (*hacked, killed, soldier*), those linking it to the attackers (*killer, muslims, nigerian*) and those linking the attack with British foreign policy (*course, foreign, policy, role*). The ambiguity of the term here – used to refer to one actor who is a *soldier* with one nationality identified and others who are murderers who also had two nationalities linked to them – perhaps begins to point towards a discourse of racism. The discourse of racism surrounding the attack will be discussed in more depth shortly. Before leaving the discussion of *british*, however, it should be noted that the identification of British foreign policy as an issue in the collocates is amplified by 16 users tweeting the title of an article, ‘Woolwich attack: of course British foreign policy had a role’, by former solider Joe Glenton in *The Guardian*.

One word which the two share in the reaction category is *muslim*. It is interesting to note that in its plural form this keyword links strongly to the murderers, yet in its singular form it links to the reaction of other Muslims, particularly the Muslim Council of Britain, to the incident. These reactions seek to distance Muslims in general from the action of Adebolajo and Adebawale. Similarly, the word *islam* reflects the distancing in the press of the religion of Islam from the actions of two followers of the religion. In the press, the keyword *islam* is associated with collocates such as *betrayal* and *condemn*. The latter collocate arises from an oft-repeated quote from the Muslim Council of Britain saying that the murder was a ‘Truly barbaric act that has no basis in Islam – we condemn this unreservedly’.

Another of the collocates of *islam, radical*, highlights what the press sees as a problem: radical Islam. This serves to distance the murderers from what one might term ‘mainstream Islam’. Another collocate of *islam* in the AP corpus is *converted*, used to refer to the process of radicalisation. Thus, the press frequently discuss the story of the conversion to Islam of the two murderers and their subsequent radicalisation. The top five collocates of *radical* in the AP are *cleric, preacher, anjem, islam* and *islamist*. In a word-sketch of *radical*, all of these collocates bar *anjem* are shown to be modified by *radical* when collocating with it. This shows the important role that the press believes so-called radical preachers and ideology have to play in discourses surrounding radicalisation. Related to this point is the other keyword we categorised as a ‘problem’ in the AP, *choudary* (referring to Anjem Choudary) – which also collocates with *radical*. In the next section, we will explore in more depth the role that Choudary has to play in the discourse surrounding the Lee Rigby attack, but for now it is suffice to say that he represents the personification of a process of radicalisation which the press argues is behind attacks such as the Woolwich attack. The warrant presented in the press for the focus on Choudary is the role he played through preaching – Adebolajo was known to listen to Choudary’s speeches and was pictured together with him at a protest. In addition, it is reported that Choudary has been linked to another person, Richard Dart, who planned to undertake a terrorist attack on British soldiers.

A final but important difference between Twitter and press reaction to the attack is coded in the political category – the press focus upon the comments of British leader David Cameron, making *cameron* a keyword. Twitter, on the other hand, yields the keyword *edl*. This word also appears frequently in the AR, which could be interpreted as a tweet about the English Defence League (EDL) being retweeted to confirm or propagate the ideological position from that tweet. Further investigation reveals a slightly different

picture, however. Particularly prominent in the Twitter data was a retweet from the official EDL Twitter account, which reads,

@Official_EDL: EDL leader Tommy Robinson on way to Woolwich now, Take to the streets
peeps ENOUGH IS ENOUGH

The EDL is ‘a movement that grew out of the football hooligan fraternity’ (Allen, 2011: 293) that espouses ‘an ideology that can be seen as a form of “new racism” or cultural racism’ (Allen, 2011) and is ‘clearly Islamophobic’ (Allen, 2011: 294). Following the Woolwich attack, its ex-leader, Tommy Robinson, staged a small-scale protest in Woolwich in response to claims that the attack was motivated by Islamic extremism.⁵ Yet, while the retweet may appear to be an endorsement of this ideology, concordancing reveals that the retweet was often a subversion of that position, as many people modified the above tweet in critique of the proposed protest. The following example shows this well – the edited part of the retweet is underlined:

@Official_EDL: EDL leader Tommy Robinson on way to Woolwich now, Take to the streets
peeps < Not in my name, cunts.

This difference between the tweets and the retweets is also observable when the collocates of the word *edl* are considered. In a word-sketch of *edl* in the AT corpus, strong collocates modified by *edl* can be divided into those which indicate a positive discourse prosody (*folk*), a neutral discourse prosody (*gathering, march, member*) and a negative discourse prosody (*thug, rioting, idiot*). These prosodies clearly indicate three quite different stances and arguably are evidence that Twitter provides a forum within which oppositional discourses co-exist. This is apparent in the word-sketch when looking at verbs that collocate with *edl* in the object position – there are ones indicating a positive discourse prosody (*support, join*) and a negative discourse prosody (*fuck, hate*). The final strong collocates modifying *edl* link it to another right wing political party – the British National Party (*bnp*). The evaluation of this party depends upon the political stance of the tweeter, once again reinforcing the notion of there being opposed discourse communities in the data. Yet by contrast, the editing of the retweets renders the retweets about the EDL as more negative. The collocate *bnp* is lost. In the word-sketch of retweets about the EDL, the positive collocates modified by *edl* remain at just one, *unite*, yet the negative collocates modified by *edl* are more numerous (e.g. *thug, knob, scum, cunt*). One new collocate being modified by *edl* in this context is *robinson*, which refers directly to the ex-leader of the EDL, with much of the abuse in the collocates described linking, ultimately, to him. At the very least, this indicates that, measured by retweeting, the discourse community opposed to the ideology of the BNP and EDL is much more successful at propagating their ideology through retweets, as evidenced by the markedly negative discourse prosody.

Another frequent set of tweets and retweets related to the EDL’s attempts to raise money for the *Help for Heroes* charity with the 4th–8th strongest collocates of *edl* (*rejects, cash, heroes, for, help*) clearly reflecting that. As indicated by these collocates, a frequent topic of conversation was the rejection by *Help for Heroes* of any money

Table 4. Usernames mentioned in tweets that achieved over 1000 retweets in the Attack Retweet corpus.

Condolence	Mrjakehumphrey (4545) HONOUROURFORCES (5015) Piersmorgan (3567, 1388) exSASChrisRyan (4973)
Arguing against a reaction against Muslims in general	MrRae1000 (4291) BBCBreaking (1462) hdck1000 (9165) SoYouWannaPwn (12,727) metpoliceuk (12,727)
Explicit evaluation of features of the attack	Piersmorgan (3212)
Muslim reaction to the murder	alialsaeed (6761) BBCBreaking (4801)
Objective report of news	BBCBreaking (2312, 2080, 1167)

donated by the EDL on the grounds that they do not accept donations from political organisations:

Woolwich: Help For Heroes Rejects EDL Cash <http://t.co/poPN15vKy2>

Discussion on Twitter was concerned with the role of the EDL in the Woolwich case. Mainly, it appeared that the apparent move by the EDL to capitalise on the religious nature of the attacks was of key concern, and most argued against the EDL-specific politicisation of the event.

While it is clear from the retweets that the anti-EDL discourse community was more active in retweeting, there are still many tweets within the AT corpus which demonstrate the racist views that this oppositional discourse was reacting against, for example,

RIP to Lee Rigby. Death and more to the coon muslim cunts who done it.

The international nature of such discourse communities is also apparent in racist tweets being contributed to the discourse by groups outside of the United Kingdom, such as the following from an American group:

Exclusive: NIGGER with bloodied hands at Woolwich (LONDON) scene ~ 14 Words Global Network14words <http://t.co/CcfM0k1M6n>

The use of retweets is clearly important in this analysis – so what may the AR corpus show us about the nature of retweeting in the Attack corpora in general? Table 4 shows the usernames appearing in tweets retweeted more than 1000 times in the AR, with the main retweets in which they were mentioned divided into discourse themes. In each case, a number cited indicates the number of times a specific tweet mentioning that user was retweeted. The tweets are organised into discourse themes in the same way as the keywords are in this article.

Of the top 10 users mentioned in the AR corpus, two are linked to news agencies – the BBC and the journalist/commentator Piers Morgan. Overall, the BBC is the most retweeted user in this corpus, with 24,413 retweets in total. The BBC retweets are generally linked to providing factual reporting from the scene of the crime, although one BBC retweet, noting the condemnation of the attack by the Muslim Council of Britain, provides Muslim reaction to the event. Another, discussing the rejection by *Help the Heroes* of funds raised for it by the EDL, at least has the effect of seeking to avoid a backlash against the Muslim community as there is an implicit rejection of the EDL and its values in the reporting of this story. Piers Morgan, while associated with the news media, is clearly responding to the event in a different way than the BBC – he expresses personal condolences at the death and tweets an evaluation of the incident, praising the courage of some of the women at the scene of the attack. In the BBC and Morgan, we see the contrast between factual journalism and the reporting of a media commentator.

Further Muslim reaction to the event is present in the form of a tweet by alialsaeed:

Woolwich murderer says he was sick of the British army killing Muslims. Well, we, Muslims, are sick of people like u using our religion to kill.

The Muslim reaction that is retweeted is clearly in harmony with the tweets that are sent which seek to work against an anti-Muslim backlash, which are typically focused on strengthening the idea that the events in Woolwich were not typical of Muslims. The most interesting tweet in this category, however, comes from SoYouWannaPwn, who posted the following tweet:

Half the things people are tweeting should put them in jail @metpoliceuk The news hasn't even got confirmed stories yet!

This tweet explains the appearance of the metpoliceuk in the list of most frequently mentioned users – this is not a tweet by them, it is a tweet aimed at them by a user who is encouraging them to intervene to stop users on Twitter from tweeting what this user believes are illegal statements. Given the frequency of the retweet, this is clearly a salient message arising from the AR corpus and one in which Twitter users are directing a message not simply at one another, but at authority.

Reaction to the conviction of Lee Rigby's murderers, 19–25 December 2013

Table 5 gives a summary analysis of the discourse in the Conviction Tweet (CT) and Conviction Press (CP) corpora following the conviction of the killers of Lee Rigby.

In the Conviction corpora, some of the themes established in the Attack corpora repeated – there is a focus on the perpetrators, the victim, the incident, the location of the incident and reaction to the incident from the Muslim community in the United Kingdom in particular. However, while there are similarities in these themes, there are also some differences worthy of note.

Table 5. The top 20 keywords in the CT and CP corpora organised by discourse theme.

Category	Tweet (CT) keywords	Press (CP) keywords
Analysis		<i>bbc</i>
Identity of the perpetrators	<i>adebolajo, adebowale, british, killers, men, michael, two</i>	<i>allah, adebolajo, adebowale, killers, michael</i>
Identity of the victim	<i>rigby, lee, soldier, british</i>	<i>afghanistan, fusilier, lee, rigby, soldier</i>
Incident	<i>murder, murdering</i>	<i>murder</i>
Location of incident	<i>woolwich</i>	<i>woolwich</i>
Motive		<i>muslims</i>
Problem	<i>muslim</i>	<i>anjem, choudary, extremist, islam, mi5</i>
Reaction		<i>muslim</i>
Verdict	<i>found, guilt, jury, trial</i>	

CT: Conviction Tweet; CP: Conviction Press.

The top five most frequent hashtags in the CT can largely be interpreted through these categories also, with a focus upon location (#woolwich, #london) and the victim (#leerigby), although there is also condolence expressed via #ripleerigby and a focus on tagging some of the tweets as #news.

First, in terms of identity, in both the attack and conviction coverage, there is a process of ‘categorisation’ (Van Leeuwen, 2008) through which Lee Rigby is identified as a *soldier*. This is more marked in the press, where the type of soldier he was is also key: *fusilier*. In the corpora, both *soldier* and *fusilier* represent Lee Rigby in terms of a particular shared identity and/or assemblage of social functions, which also represent relationships with wider ideologies and social institutions, as did *drummer* and *soldier* in the earlier initial reaction corpora. On the other hand, Adebowale and Adebolajo are categorised in the earlier Attack corpora by their status in law (*suspect* in Twitter) and religion (*muslims* both in Twitter and in the press), while in the Conviction corpora, the focus is upon their actions as *killers* (both in Twitter and in the press). And in both of the Conviction corpora, their act of *murder* is key. The switch away from highlighting their religious affiliation in their categorisation is arguably congruent with a discourse in the previously discussed Attack corpora which sought to set them aside as non-mainstream followers of Islam. Now, in the Conviction corpora in particular, the focus of the label *soldier*, which is not upon the actions of soldiers (the only keyword linking Lee Rigby to an action is *drummer* in the earlier Attack corpora), also allows a contrast to develop between the victim, who is not explicitly linked to acts of violence, and the killers, who are categorised in the Conviction corpora very clearly by the act they carried out when they attacked Lee Rigby (*murder, murdering*). It also provides a disjunct between the identity of the victim and the publicly expressed motivations of the killers themselves, as widely reported in the AP corpus to the extent that the keyword *muslims* is linked with this quote, as in the following extract from the corpus:

In a chilling rant captured on camera, the knife-wielding man declared: ‘The only reason we have killed this man today is because Muslims are dying daily by British soldiers’.

A further way in which the religious motives of the killers are still present in the press is clear in the keyword *allah*. This word becomes key in the CP corpus through frequent reporting of the murderers' claims either to be acting in the name of or to be guided by *allah*. This is clear in the top collocates of *allah* in the CP – which include *command*, *decree* and *obeying*. Similarly in a word-sketch of *allah*, the top two collocates modified by 'of Allah' are *soldier* and *command*. These collocates largely derive from self-representation by the killers as reported in the press:

Adebolajo – who described himself as a 'soldier of Allah' – also said ... 'We decided to wait in the vicinity of the barracks that is in Woolwich', ... 'By the command of Allah, by Allah's decree.'

While Allah clearly indirectly references the religion of Islam, the representation of the killers does not root their attack clearly in the religion as a whole; rather, it simply suggests that their inspiration was personal in relation to the God of that religion. This again works overall with a discourse which seeks to marginalise the acts of the murderers in the context of the religion itself. The acts appear personally inspired. They are not condoned by the religion, although the murderers claim it is the plight of Muslims as well as the will of Allah which moved them to act. Reinforcing the view that their acts are not condoned by fellow Muslims is the keyword *muslim* in the reaction category for CP. It is related to the prominence given to reporting of British Muslims distancing themselves and their religion from the acts of the murderers, as in this example of press reaction to the conviction:

numerous British Muslims already strongly condemn Islamist violence. The Muslim Council of Britain, for example, called the murder of Lee Rigby 'truly a barbaric act'.

Unsurprisingly, the judicial process and the verdict feature in the conviction corpus. However, it is Twitter which reports the judicial process most saliently. This almost certainly links to the immediacy of Twitter reaction – the verdict as 'breaking news' is picked up proportionately more on Twitter, compared to the press. Another trend visible in the Attack corpora is also visible in the Conviction Corpora – the press data are more clearly linked to analysis and the search for explanations than the Twitter data. Once again, extremists are identified as a problem, as is the version of Islam that they espouse. The focus upon one extremist, Anjem Choudary, one-time leader of the banned organisation al-Muhajiroun, is once again marked. Choudary is clearly framed as a problem by the press, particularly because of his public refusal to condemn the attack on Lee Rigby, in a way that he is not in the Tweet corpora, revealing this fundamental difference between Twitter and press reaction. By way of illustration, in the CP the top 10 collocates of Choudary (barring his first name and title) are *preacher*, *hate*, *condemn*, *give*, *refused*, *airtime*, *comments*, *interview*, 's and *former*. The word-sketch of *choudary* is also revealing. In the word-sketch, the word *choudary* is modified by *disgusted*; for example, *people should be disgusted Mr Choudary will not condemn*. Similarly, the word *influenced* is linked to *choudary* as in examples such as *he was heavily influenced by Mr Choudary's*

preaching. The following quote is illustrative of the press reaction to Choudary and the role they believe he plays in spreading violent radical Islam in the United Kingdom:

At some point in his transition from happy and innocent child to jihadist obsessed with war, Jaman fell under the spell of the fanatical British preacher, Anjem Choudary. Yesterday we heard Choudary's poisonous creed broadcast on the airwaves when he appeared on Radio 4's Today programme and repeatedly refused to condemn the barbaric murder last summer of Drummer Lee Rigby.

Choudary's refusal to condemn the murder of Lee Rigby also included a re-insertion into the discourse of the motivation of the murderers. It also implicitly highlights the role of the British armed forces in state legitimised acts of violence through war, thus furthering the distance between the discourse strategy of Choudary and the press, as in the following example from the CP:

Mr Choudary said: 'I think that to talk about condemnation or to talk about how we feel is not the most important question now, and I'm not going to go down that road. I think that what is important is to learn lessons from what has taken place. Whether you agree or disagree with what took place, you cannot predict the actions of one individual among a population of 60 million when the Government is clearly at war in Muslim countries. I condemn those who have caused what has taken place on the streets of London, and I believe that the cause of this is David Cameron and his foreign policy'. Reacting to Mr Choudary's comments the Muslim Council of Britain said the BBC was 'unwise' to give him airtime. A spokesman said their condemnation of Lee Rigby's murder comes from a 'deeply held belief in Islam and the premium it places on the right to life'.

The juxtaposition of the views of Choudary in this example with those of the Muslim Council of Britain has the overall effect of representing him as a dangerous fringe figure who is not in the mainstream of Islam. This is entirely consonant with the discussion of the keywords in the 'problem' category in both the Attack and Conviction corpora. Choudary is discussed, and in similar terms, in the Twitter data, although the discussion focuses on his media appearance, especially a BBC Radio 4 interview he gave. It also, however, manifests itself in the keyword *muslim* in the CT being linked to the problem category through Choudary. The salience of the word *muslim* in the Twitter reaction to the conviction is due to the widespread reporting of the views of Choudary, who is described as a *muslim leader* (*war*, *warns* and *leader* are the three strongest collocates of *muslim* in the CT). His views are widely condemned, as in the following example, where a Tweet from Sky News is quoted and responded to in a retweet. The response is underlined in the example below:

@SkyNews: Woolwich Murder: Muslim Leader Warns Of 'War' ... < Why would you give publicity to stupidity?

This retweet was then retweeted again with a further reply sent to Choudary's Twitter address saying simply, @anjemchoudary How does 'fuck off' sound?

Table 6. Usernames mentioned in tweets that achieved over 1000 retweets in the Conviction Retweet corpus.

Family Reframing	OfficialLRigby (2721) Its_UsY (2869) MMLdn (1440)
News	BBCBreaking (1100)

It is interesting, nonetheless, to see that the case of Choudary shows a clear link and interaction between social and traditional media.

Before leaving the discussion of the problem category, however, the appearance of the keyword *mi5* in this category bears some discussion. MI5 in particular is reported as a problem in the CP – MI5's attempt to recruit one of the murderers and their subsequent failure to act in a way that would stop the murder occurring is the subject of widespread criticism in the corpus.

Moving to consider the Conviction Retweet (CR) corpus, Table 6 looks at tweets which received over 1000 retweets in the CR.

Setting aside the retweet from BBCBreaking, announcing the verdict, which was heavily retweeted, two themes come across very strongly in this table. First, the sympathy that was present in the AT and AR corpora is still present in the CR corpus – a very popular retweet is from the user OfficialLRigby, which is a picture of Lee Rigby with his son Jack. The other theme in the retweets is linked to reframing the incident in terms of racism. Like condolence, this was present in the AT analysis but not in the CT analysis. A popular retweet from MMLdn asks readers to look at the reporting of an attack by white racist murderers and question why the story had not received the same coverage as the Woolwich killing: ‘Do they look black or muslim?’ the retweet asks. As noted, this picks up on some of the racist themes arising from the reaction to the attack – it implies that the attack is getting attention because the attackers were black and Muslim and, presumably, because the victim was white and non-Muslim. This is an interesting example of how retweets can in themselves represent an invitation to reframe an event – the strong implied endorsement of this view through retweeting indicating that there is sympathy for this point of view from some Twitter users, consequently making the argument more salient than it is in the CT corpus.

A very similar theme is present in the retweeting of a tweet from ItsUsY relating to the same story as the MMLdn tweet, asking ‘Why hasn’t this got the same coverage as the Woolwich thing?’. This is an example of *thing* being used to encapsulate the whole event, arguably including the discourse surrounding the events under discussion. The tweet was trying once again to reframe the discussion around Woolwich in terms of race and religion. The construct ‘the Woolwich thing’ may be said to be reflective of the author’s attitude to the event – looking in a large general English corpus⁶ for collocates of ‘the X thing’ reveals that a dismissive (e.g. ‘I was happy to laugh the whole thing off’), disapproving (‘the whole thing is loaded with dread’) or disinterested (‘I guess the whole thing was a kind of bother’) attitude seems to be associated with this construction.

Table 7. The top 20 keywords in the ST and SP corpora organised by discourse theme.

Category	Twitter (ST) keywords	Press (SP) keywords
Criminal		
Identity of the perpetrators	<i>adebolajo, adebowale, killers, michael, british, murderers</i>	<i>dennehy</i> <i>adebolajo, adebowale, killers, allah, michael</i>
Identity of the victim	<i>rigby, lee, soldier, british</i>	<i>rigby, lee, solider, fusilier</i>
Incident	<i>murder</i>	<i>murder, killing</i>
Location of incident	<i>woolwich</i>	<i>woolwich</i>
Justice system		<i>judge, justice, bailey, dock</i>
Problem		<i>islam</i>
Reaction	<i>justice</i>	
Sentence	<i>sentenced, prison, sentence, term, jail, jailed, sentences, sentencing</i>	<i>whole-life, sentencing</i>

ST: Sentencing Tweet; SP: Sentencing Press.

The five most frequent hashtags in the ST corpus focus on the location (#woolwich) and the identity of the victim (#leerigby, #rigby). As with the Conviction Tweet corpus, condolence is also expressed through hashtags (#ripleerigby, #lestweforget).

Reaction to the sentencing of Lee Rigby's murderers, 26 February–4 March 2014

Table 7 gives a summary analysis of the top 20 keywords in the Sentencing Tweet (ST) corpus and the Sentencing Press (SP) corpus gathered following sentencing.

The most striking feature of the analysis of the ST and SP corpora is the stability of the discourse from conviction to subsequent sentencing. Some themes have altered – while the keyword *islam* is still associated with the perceived problem of radical Islam in the press, the personalising of that issue to Anjem Choudary has abated. The question of why is of interest – there is little doubt that this, in part at least, derives from his media profile diminishing. In the SP corpus, there are only five mentions of Anjem Choudary (3.6 per 100,000 words), none of which are associated with either direct or indirect quotations from him. Contrast this with the 211.9 mentions per 100,000 words in the CP corpus and the prevalence of words collocating with *choudary* in the CP indicating direct or indirect reporting of his views – such as *opinions, comments, speak, interview* and *said*. Choudary's prominence in the media in the earlier two press corpora almost vanishes in the SP corpus, despite the fact that the prevailing view that he had a connection to and may have inspired the killers has not changed as a claim, as the following example from the SP corpus shows. This example follows a lengthy quote from a speech Adebowale gave at an event. The speech seemed to foreshadow the sentiments he expressed after killing Lee Rigby:

Choudary denies meeting Adebowale there or organising the event – although his mobile number is on the web poster.

A keyword found in the SP corpus which was not present in the Conviction corpora was *whole-life*, caused by the whole-life term given to Adebolajo. This was reported

widely because the whole-life term is rarely used.⁷ Another SP-only keyword relates to this as *dennehy* refers to another murderer being sentenced in the same court as the Woolwich killers later in the same week.⁸ The previous trial is mentioned in the context of the reporting of Dennehy's trial, producing *dennehy* as a keyword. The mentioning of the Woolwich killers in the context of the Dennehy reporting, however, also serves to bracket the Woolwich killers, who were ideologically inspired, with Dennehy, who was a sadistic serial murderer.

Note that the word *justice* is categorised differently between the Sentencing Twitter and Press corpora. This is because in the press the word tends to be a title afforded to the judge – Mr Justice Sweeney – who presided in the Woolwich case. The papers quote his summing up extensively and refer to him by his formal title frequently. The focus on Justice Sweeney in the press is marked. What is also marked is how consonant Justice Sweeney's summing up is with the press discourse at the time of the attack and conviction. The following quote from Justice Sweeney in the SP corpus could almost have been used to paraphrase the discussion of the CP corpus earlier in this article:

Mr Justice Sweeney said: 'You each converted to Islam some years ago. Thereafter you were radicalised and each became an extremist, espousing views which, as has been said elsewhere, are a betrayal of Islam'.

Justice Sweeney and his views are not at all prominent on Twitter, however. Twitter users focus relatively more on the concept of *justice*, the top six collocates of this keyword being *served*, *has*, *been*, *real*, *done* and *for*. Partly these come from tweeting of a quote from Lee Rigby's family saying that they are 'satisfied that justice has been served' in the case (this phrase occurs in 30 tweets in the STs). Yet this is not the only use of this word – the notion of whether justice was served is both contested and politicised. It is contested in tweets such as

the 2 males who murdered Lee Rigby received life sentences, 1 actual life and 1 minimum of 45 yrs has justice been served?

The collocate *for* is closely linked to the sixth strongest collocate for *justice* in the ST corpus – *real*. Both arise from tweets of what appears to be a syndicated news story tweeted by various news sources in which the phrase 'BNP leads the call for REAL Justice for Lee Rigby' is used (this phrase occurs in 119 tweets). So, while many elements of the discourse between the conviction and the sentencing are stable, there are differences, with some elements falling away (Choudary) which had been salient up to that point, some new elements (notably a discussion of whether justice had been served in the ST corpus), and a theme from the attack reasserting itself through the attempts of a right wing party, the BNP, to politicise the event. Yet the discourse community seeming to wish to bring the BNP back to salience is the press through its Twitter presence. Retweeting, however, may have altered this for, as noted, it has the potential to show which of the themes identified in the ST corpus were amplified through retweeting by discourse communities on Twitter. Hence, let us consider the SR corpus. What this shows is that, despite the interest of the news agencies in the BNP, Twitter users did not retweet

the story widely. In fact, the retweeting around the sentencing is at a much lower level than for the attack or conviction. Only two tweets, both from SkyNewsBreak, pass the threshold set in this article for analysing retweets – one was a tweet announcing the sentencing of Adebowale (1355 retweets) and the other a tweet announcing the sentencing of Adebolajo (1200 retweets). So the BNP story, despite being tweeted widely by the press, does not seem to have garnered much in the way of retweets. In this case, the discourse communities on Twitter seemed resistant to this tweet, although they were not indifferent to the story, as the retweeting of the sentencing shows. This supports a view that there was not wide ideological support for the BNP move. However, another way of exploring the influence of the BNP story might be to see whether it could be linked to the SR corpus containing more retweets focusing upon claims that justice has not been served in the case. Given that (1) there were tweets which said this and (2) this was the thrust of the BNP's proposed campaign, then if the SR corpus contains a stronger focus on tweets claiming justice had not been served, it might be argued that the BNP campaign influenced the discourse or simply that those retweeting felt that justice had not been served. Yet, when the collocates of *justice* are explored in the SR, it is clearly the quote from the family of Lee Rigby that is dominant as the top five collocates are *served*, *done*, *has*, *been* and *satisfied*. The collocate *real* associated with the word *justice* in the ST corpus has fallen away from the top five collocates (it is the 51st strongest collocate of *justice* in the SR corpus), indicating that the story was relatively dispreferred in retweeting.

Conclusion

Overall, this article has demonstrated that using CL it is possible to approach large bodies of text and begin to draw out the key elements in the discourses contained in those texts. This is true of short texts (tweets) as well as longer texts (newspaper articles). While some of what is found may be obvious – the focus on the place where the attack took place, the perpetrators and the victim, for example – other more subtle themes emerged in the analysis (e.g. attempts to distance the murderers from the religion of Islam). In this conclusion, we will focus on three key areas where we think the scale of the corpus analysis of these data is of assistance in looking at a complex event like this which stretches over time. The three features we would point to are the relationship of the press to Twitter, the attribution of blame and the search for explanations, and the role of retweets in analyses of Twitter data.

There is little doubt that if we look at a major news story like the Woolwich attack, then the presence of news sources on Twitter, especially in the form of retweets, will be dominant. For example, if we look at the top 10 retweeted users in the SR corpus, seven are news sources (bbcbreaking, skymarkwhite, skynewsbreak, itvnews, courtnewsuk, thesunnewspaper and bbcnews). Yet the line drawn in this article between Twitter and the press is highly porous. Twitter users do not ignore the press – indeed many of the most retweeted users in the Retweet corpus are press accounts. However, the porosity is largely one way – in this case study, as noted, there is a substantial traditional media presence on Twitter, with non-press users retweeting links to press stories frequently. On the other hand, for this news story, the press does not seem to respond directly to Twitter

despite the role that social media played in initial reporting of the attack. The one possible exception is the press reportage of the EDL. However, what the press reports is the protest staged by the EDL and the rejection of money they gathered by *Help the Heroes*. It is not a discussion of what is being said on Twitter – the press appears largely deaf to the attempts to reframe the discussion on Twitter, for example. So while porous, the direction of influence seems to be from the press to Twitter. With that said, it would be wrong to give the impression that Twitter users simply retweet press content indiscriminately. The case of the BNP ‘REAL justice’ tweet is a good case in point and provides evidence for our position that retweeting represents an ideological choice on behalf of Twitter users – the press heavily tweeted the BNP story, but it was not widely taken up and amplified by Twitter users through retweeting. As well as ignoring press messages, Twitter users are also willing to use their fellow Twitter users as sources of news, with the frequent tweets in the wake of the attack on Lee Rigby revealing, through the keyword *happened* (and associated forms), that users were turning to others on social media to find out what was happening.

One highly salient difference between the press and Twitter, however, relates to the attribution of blame and the search for explanations. That is much more marked in the press, where a very clear narrative develops, and is adhered to, that distances Islam from the killers and identifies them with an extremist form, a ‘betrayal’, of the religion that is not typical of that practised by the majority of Muslims. The press provide warrants in support of this view from the Muslim Council of Britain, and personalise the discussion around Anjem Choudary who is constructed in very negative terms. Evidence for this being the mainstream public discourse surrounding the event is provided by the summing up in the case, where the judge characterises the attack in strikingly similar terms. The discourse also sums up the response to an ideologically inspired murder – the ideology is marginalised, the killers are divorced from the religion associated with the ideology, the ideologue taken as representative of the ideology is the focus for much public anger and the ideology is cast out of the religion not merely by the press but by relevant religious authorities. This response is echoed on Twitter, but the reaction to the event itself, personal sympathy for the bereaved and the clash between the British identity of the killers and their acts seem to be stronger pre-occupations of Twitter users. Some Twitter users do try to reframe the discussion in general terms relating to racism, but that discourse fades across the timeline of the event until it is absent in the sentencing corpora.

Retweeting represents a way into exploring discourse communities and ideological positioning on Twitter. As has been shown, retweeting can be important in showing how oppositional discourses on Twitter are perceived and either promoted or ignored by users. Similarly, retweeting is an important source of commentary, especially where retweeting of modified tweets is concerned. Also, retweets may be used to amplify messages directed at authority on Twitter, as the example of the tweet exhorting the Metropolitan Police to take action against other Twitter users showed.

If a criticism of the corpus approach in particular, and perhaps discourse analysis in general, is that it focuses too often on the text itself, looking less at processes of production and reception, then the use of Twitter data in particular begins to counter that. The mode of production of Twitter data has been essential to the interpretation of some of the analyses in this article, with the dialogic nature of Twitter being a key to some of what has

been said. Importantly, retweets not only give further weight to the observation about the dialogic nature of Twitter, they are also an important proxy for reception – tweets evoke a series of responses, all of which can be, to a lesser or greater degree, inferred from the retweet corpora. Some tweets are ignored. Some are retweeted, some with critical commentary or an endorsement attached. Retweets can in turn be retweeted, with the process of critical commentary becoming nested. In short, the discourse communities on Twitter can use retweeting as a way of encoding reader reception in the corpus itself. This was shown in the analysis of the AT and AR corpora in particular, where a contested discourse around racism in the AT corpus seems to be resolved firmly in favour of an anti-racist discourse in the AR corpus. Given that this is happening in short texts, the impulse to comment and invest time in editing messages to fit in commentary is notable. As such, our claim that retweets may be seen as a proxy for reception is well made in our view.

Funding

The research presented in this article was supported by the Economic and Social Research Council (ESRC) Centre for Corpus Approaches to Social Science, ESRC grant reference ES/K002155/1

Notes

1. There is a low level of noise in the corpora, for example, tweets including the word *woolwich* which are not about the Woolwich attack. For example, one of the most frequent sources of error in the corpus of tweets gathered in the week of the attack are 13 tweets in the corpus which refer to jobs available in the Woolwich area. However, such noise in the data is very small in comparison with the size of the dataset (419,775 tweets in the case of the Twitter attack corpus).
2. See <http://www.laurenceanthony.net/software.html>
3. See <https://www.sketchengine.co.uk/>
4. See the Analytical approach section for details of how this was done.
5. The English Defence League (EDL) is clearly evaluated as racist by some Twitter users who attach the hashtag #racism (the third most frequent hashtag in the Attack Tweet (AT) corpus) to EDL (*edl* is the fifth strongest collocate of #racism).
6. In this case, the British National Corpus.
7. At the time of writing, the whole-life term has been imposed 23 times in the United Kingdom since the 1950s.
8. Given that Dennehy was also given a whole-life term, it must be noted that the salience of this keyword is amplified accordingly. In the Sentencing Press (SP) corpus, the keyword *whole-life* collocates with *adebolajo* 50 times, while it collocates with *dennehy* only seven times; hence, the interpretation of the keyword *whole-life* as relating principally to *adebolajo* seems justified.

References

- Allen C (2011) Opposing Islamification or promoting Islamophobia? Understanding the English Defence League. *Patterns of Prejudice* 45(4): 279–294.
- Bachmann I (2011) Civil partnership – ‘gay marriage in all but name’: A corpus-driven analysis of discourses of same-sex relationships in the UK Parliament. *Corpora* 6(1): 77–105.
- Baker P (2004) ‘Unnatural acts’ Discourses of homosexuality within the House of Lords debates on gay male law reform. *Sociolinguistics* 8(1): 88–106.

- Baker P (2006) *Using Corpora in Discourse Analysis*. London: Continuum.
- Baker P (2010) Representations of Islam in British broadsheet and tabloid newspapers 1999–2005. *Journal of Language and Politics* 9(2): 310–338.
- Baker P (2014) *Using Corpora to Analyze Gender*. London: Bloomsbury Publishing.
- Baker P and McEnery T (2005) A corpus-based approach to discourses of refugees and asylum seekers in UN and newspaper texts. *Journal of Language and Politics* 4(2): 197–226.
- Baker P and McEnery T (in press) Who benefits when discourse gets democratised? Analysing a Twitter corpus around the British Benefits Street debate. In: Baker P and McEnery T (eds) *Corpora and Discourse Studies: Integrating Discourse and Corpora*. Basingstoke: Palgrave.
- Baker P, Gabrielatos C and McEnery T (2013) *Discourse Analysis and Media Attitudes*. Cambridge: Cambridge University Press.
- Baker P, Gabrielatos C, KhosraviNik M, et al. (2008) A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society* 19(3): 273–306.
- Gabrielatos C and Baker P (2008) Fleeing, sneaking, flooding: A corpus analysis of discursive constructions of refugees and asylum seekers in the UK Press, 1996–2005. *Journal of English Linguistics* 36(1): 5–38.
- LexisNexis (2014) Lexis news search. Available at: <http://www.lexisnexis.com/uk/nexis/search/loadForm.do?formID=GB01NBSimplSrch&;random0.7450459500123497> (accessed 15 April 2014).
- McEnery T, Xiao R and Tono Y (2006) *Corpus-based Language Studies*. Abingdon: Routledge.
- Mautner G (2009) Corpora and critical discourse analysis. In: Baker P (ed.) *Contemporary Corpus Linguistics*. London: Continuum, pp. 32–46.
- Partington A, Morley J and Haarman L (eds) (2004) *Corpora and Discourse*. Bern: Peter Lang.
- Rychlý P (2008) A lexicographer-friendly association score. In: *Proceedings of Recent Advances in Slavonic Natural Language Processing*, RASLAN, Karlova Studánka, 5–7 December, pp. 6–9. Brno: Masaryk University.
- Stubbs M (1996) *Text and Corpus Analysis*. London: Blackwell.
- Stubbs M (2001) *Words and Phrases: Corpus Studies of Lexical Semantics*. London: Blackwell.
- Van Leeuwen T (2008) *Discourse and Practice: New Tools for Critical Discourse Analysis*. London: Routledge.

Author biographies

Tony McEnery is Distinguished Professor of English Language and Linguistics at Lancaster University. His research involves applications of corpus linguistics, particularly in the social sciences. He is currently Director of the ‘Corpus Approaches to Social Science’ Research Centre (CASS) at Lancaster University. CASS is funded by the United Kingdom’s Economic and Social Research Council. Professor McEnery’s latest book is *Corpus Linguistics: Method Theory and Practice* (Cambridge University Press, 2012), with Andrew Hardie.

Mark McGlashan is a Senior Research Associate at CASS in the Department of Linguistics and English Language at Lancaster University, where he is also a PhD candidate in Applied Linguistics. His recent research focuses on the application of corpus methods to the detection and analysis of abusive and/or extremist online discourse communities largely in the context of Twitter. His research interests include social inequalities, Critical Discourse Analysis, Corpus Linguistics, and Social Network Analysis.

Robbie Love is a PhD candidate in Linguistics at CASS in the Department of Linguistics and English Language at Lancaster University. His research interests include spoken corpus compilation, corpus-based critical discourse analysis and language change.