Sketching Muslims: A Corpus Driven Analysis of Representations Around the Word 'Muslim' in the British Press 1998–2009

[1,]*PAUL BAKER, [2]COSTAS GABRIELATOS and [1]TONY McENERY

[1]Department of Linguistics and English Language, Lancaster University and [2]Edge Hill University
*E-mail: p.baker@lancaster.ac.uk

This article uses methods from corpus linguistics and critical discourse analysis to examine patterns of representation around the word *Muslim* in a 143 million word corpus of British newspaper articles published between 1998 and 2009. Using the analysis tool Sketch Engine, an analysis of noun collocates of *Muslim* found that the following categories (in order of frequency) were referenced: ethnic/national identity, characterizing/differentiating attributes, conflict, culture, religion, and group/organizations. The 'conflict' category was found to be particularly lexically rich, containing many word types. It was also implicitly indexed in the other categories. Following this, an analysis of the two most frequent collocate pairs: *Muslim world* and *Muslim community* showed that they were used to collectivize Muslims, both emphasizing their sameness to each other and their difference to 'The West'. Muslims were also represented as easily offended, alienated, and in conflict with non-Muslims. The analysis additionally considered legitimation strategies that enabled editors to print more controversial representations, and concludes with a discussion of researcher bias and an extended notion of audience through online social networks.

INTRODUCTION

On 18 October 2006, the British tabloid[1] newspaper *The Daily Star* carried the front page headline 'BBC PUT MUSLIMS BEFORE YOU!' The headline was accompanied by a picture of a woman in a face-covering niqab making a two-fingered gesture. Despite the fact that the 2001 Census recorded 1,588,890 Muslims living in the UK, comprising 2.8% of the population,[2] the use of the second person pronoun *you* to directly address the reader seems to exclude the possibility that a Muslim could buy the newspaper or even read the headline. During the first decade of the 21st century, and particularly since the 9/11 attacks on the USA, British newspapers have played an increasingly important role in encouraging and influencing the direction of a national debate about Muslims and Islam. Between 2000 and 2009,

the word *Muslim* and its plural appeared 121,125 times in the national British press (∼33 times a day).

The analysis in the article is a small, but salient, part of an ESRC funded project on the representation of Islam and Muslims in the British press.[3] We explored representations of words like *Muslim*, *Islam*, and *Islamic* and examined how a wide range of topics relating to Islam were covered in the media, including debates about veiling, the threat of extremism, and Muslims receiving state benefits, as well as considering change over time and variation between different newspapers. The analysis of such issues allowed us to more widely consider different representations of Muslims beyond the word *Muslim* itself, for example, by letting us focus on frequently named individuals. This article only focuses on the ways that the press has constructed the word *Muslim* during the period under study, by using methods associated with corpus linguistics and critical discourse analysis (CDA).

However, in an age of multiple digital television channels, some exclusively devoted to news, and a move towards online news consumption through political blogs and sites like Twitter, it could be argued that the influence of traditional newspapers has declined. Figure 1 shows how sales figures have generally fallen between 2000 and 2009.

The bestselling national newspaper, *The Sun*, sold only 2,899,310 copies in December 2008.[4] Thus, it could be argued that *Sun* readers represent a small percentage of the British population that stood at 57 million in the official Census of 2001.

Figure 1: Newspaper sales October 2000–December 2010 (daily newspapers) through Audit Bureau of Calculations. Figures derived from http://www .guardian.co.uk/media/abcs

The number of newspapers sold is not the same as the number of people who read a newspaper, however. The National Readership Survey's study of 36,000 adults estimated that ∼7,860,000 people read *The Sun* every day during July 2008–June 2009. Conboy (2010: 2) argues that new technologies have 'forced newspapers to alter the structure and address of their language, as they bid to retain a profitable and influential share of the market...' One strategy has been a move towards tabloidization (involving increases in celebrity news along with more sensational and personal stories, and a decrease in political and international news). Additionally, most newspapers have engaged with new digital technologies, publishing online versions, which for the most part are free and reach a wider audience, including international readers. Therefore, although sales of the paper versions of newspapers may be declining, other evidence suggests that newspapers continue to be influential, hence, our interest in representations in the press.

The following section attempts to locate our study among existing research that addresses the representation of Muslims in the western media and to justify our methodological approach (combining corpus linguistics with CDA). We then describe the way in which we collected our data, and how it was analysed. This is followed by two results sections, the first takes a general 'broad-brush' approach that focuses on common noun collocates of the word *Muslim*. The second is a smaller-scale detailed analysis that concentrates on the two most common lexical collocates of *Muslim*: *community* and *world*. Finally, our conclusion reflects on our findings and the research process.

RESEARCH ON MEDIA LANGUAGE AND ISLAM

Research that has examined the representation of Islam in the western news media has generally found evidence for negative bias. Awass (1996) studied US media, finding that news articles linked Islam to fundamentalism and terrorism. Dunn (2001: 296), who focused on two Australian newspapers, notes how Muslims were constructed negatively: 'fanatic, intolerant, fundamentalist, misogynist [and] alien' 75% of the time, whereas positive constructions accounted for 25% of cases. Richardson (2004b) argues that negative media portrayals have pushed British Muslims to the margins of society, whereas Richardson's (2004a) analysis of British broadsheets for a 4-month period in 1997 found four common *topoi* (persuasive strategies or warrants that connect an argument to a claim or conclusion) associated with reporting of Islam: as a military threat, as terrorists/extremists, as a threat to democracy, and as sexist or a social threat. Akbarzadeh and Smith (2005: 4) argue that '...recurring language used to describe Islam and Muslims (such as ''Islamic terrorism,'' ''Muslim fanatics'') can come to be representative of all Muslims and Islam as a religion'. Poole and Richardson (2006) note that 'distorted reporting... often characterizes the coverage of Islam and Muslims'. Such findings are also in-line with observations and findings made by the government and

government agencies, most notably reports from the House of Commons Home Affairs Select Committee.[5]

As online electronic archives of newspaper articles are gradually becoming more easily available, quantitative approaches have become a more popular method of analysis, by using either content analysis or corpus linguistics procedures. As an example of the former, Moore *et al.* (2008) applied a categorization scheme to 974 articles on Islam in the British press from 2000 to 2008 and examined a smaller sample of visuals appearing in the articles. They found that stories that focused on extremism or differences between Islamic culture and 'The West' increased over time, whereas stories that focused on attacks on Muslims or problems that they faced decreased. Their visual analysis found a high use of police mugshots used to portray Muslims, a greater number of pictures of Muslim males compared with females, and a high number of pictures of Muslims engaged in religious practices.

A corpus-based approach to the representation of Islam was taken in a pilot study[6] for this article (Baker 2010) that compared 22 million words of articles about Muslims and Islam in British tabloid newspapers with 65 million words of similar articles in broadsheets. A list of keywords (words that were statistically more frequent in one set of data, when compared against each other) was derived and then subjected to more detailed analysis of context through concordances. The tabloids were found to focus more on stories that related Islam to terrorism and extremism, often writing about controversial Muslim clerics like Omar Bakri and Abu Hamza. The broadsheets were more likely to write about Muslims in a wider range of contexts, although they tended to have many more stories about Muslims involved in conflicts around the world. The study also traced how two newspapers, the *Daily Mail* and *Daily Express* referred to Muslims as *Moslems*, and how the *Mail* continued to use this term for a year after the Muslim Council of Britain asked it to stop doing so, as the spelling evoked a pronunciation that was similar to the Arabic word for oppressor.

Like our pilot study, the analysis in this article also takes a corpus-linguistics approach to investigate the representation of Islam. The use of large collections of electronically encoded data, in conjunction with computer software, presents a number of advantages for discourse analysts. Mautner (2009: 34) summarizes these as helping to give analysts an initial focus in their data, lending credence to generalizations, and counteracting the criticism that researchers within CDA supposedly 'cherry-pick' small amounts of data that support pre-conceived ideologies (see Widdowson 2004: 102). Early proponents of this approach include Hardt-Mautner (1995) and Stubbs (1996). Partington (2006) refers to the relatively new field of Corpus-Assisted Discourse Studies, and it is under this term that we locate our own research.[7] This study does not aim to replicate typical modes of CDA (e.g. analysis of transitivity or argumentation strategies), albeit using corpus data instead of a small number of selected texts. Rather, our work adopts and adapts concepts, constructs, and metrics developed within corpus linguistics to achieve aims compatible with those of CDA, in that it aims to examine social issues 'that are

caused or exacerbated by public text and talk, such as various forms of social power abuse (domination) and their resulting social inequality' (van Dijk 2009: 63; see also Bloor and Bloor 2007: 12; Wodak and Meyer 2009: 8–10).

However, our approach is also one that draws on concepts and practices used in critical discourse analysis (e.g. Fairclough 1995; Reisigl and Wodak 2001; Wodak and Meyer 2009). More specifically, we were broadly influenced by elements of Fairclough's textually oriented discourse analysis, which involves 'linguistic *description* of the language text, *interpretation* of the relationship between the (productive and interpretative) discursive processes and the text, and *explanation* of the relationship between the discursive processes and the social processes' (Fairclough 1995: 97, original emphasis). Therefore, we have examined the ways that processes of production impact on what can and cannot be written. For example, we have looked at how regulatory bodies like the Press Complaints Commission help to set boundaries for journalism. Additionally, we have considered legitimation strategies (van Leeuwen 2007) or techniques that are used to explain and justify particular representations and intertextual relationships (Reisigl and Wodak 2001: 37–41) or ways that texts refer to other texts.

DATA AND METHODS

The online newspaper database, *Nexis UK*, was used to collect newspaper articles that contained words relating to Muslims and Islam. The following search term (developed through trial and error) was used to query daily and Sunday versions of national British newspapers[8] held in the archive, between 1 January 1998 and 31 December 2009:

> alah OR allah OR ayatollah! OR burka! OR burqa! OR chador! OR fatwa! OR hejab! OR imam! OR islam! OR koran OR mecca OR medina OR mohammedan! OR moslem! OR muslim! OR mosque! OR mufti! OR mujaheddin! OR mujahedin! OR mullah! OR prophet mohammed OR q'uran OR rupoush OR rupush OR sharia OR shari'a OR shia! OR shi-ite! OR shi'ite! OR sunni! OR the prophet OR wahabi OR yashmak! AND NOT islamabad AND NOT shiatsu AND NOT sunnily

The search term was not case sensitive, and the punctuation mark,!, acted as a wildcard for any set of characters. Although the search term would not have captured articles that referred only to individual Muslims by name or oblique references, such as 'a certain religion', we feel that it elicited a large enough number of articles for us to claim considerable representativeness. In all, the corpus stands at 200,037 articles collectively consisting of 143 million words.

Our overall methodology has involved a combination of corpus-driven and corpus-based approaches, the former lets the analysis be driven by whatever is frequent or salient in the data, the latter allows users to test pre-existing hypotheses (see Tognini-Bonelli 2001). For example, we have examined

keyword and frequency lists, and we have followed up certain pre-set lines of enquiry based on issues raised by other scholars or in the pilot study (Baker 2010). For the purpose of this article, the analysis focuses on the word *Muslim*, which was frequent in the corpus data and also highly salient in terms of being the term that directly refers to the identity associated with the religion of Islam. Therefore, in terms of understanding how a set of people are constructed in the press, *Muslim* is perhaps the most relevant word to subject to a close analysis. In doing so, our general research question is 'What does a collocational analysis of the word *Muslim* reveal about the construction of this group?'

The first stage of the analysis involved obtaining a word sketch of *Muslim* using the online corpus analysis interface Sketch Engine (Kilgarriff *et al.* 2004). Sketch Engine allows corpora to be uploaded onto an online database; the corpora are then grammatically tagged in such a way that when collocates are derived, it is possible to identify grammatical relationships between collocates. For example, Pearce (2008) used Sketch Engine to derive word sketches of the lemmas MAN and WOMAN in the British National Corpus, and found that WOMAN tended to be the grammatical subject of verbs that held a prosody for annoyance, such as *annoy*, *cluck*, *fuss*, *nag*, whereas it was the grammatical object of verbs that referenced sexualization, such as *bed*, *ravish*, *shag*. Sketch Engine, therefore, provides a more sophisticated picture of collocational patterns than merely considering pairs of words together.

A word sketch is necessarily a somewhat 'broad-brush' approach, and we have argued in earlier research focused on representation of immigration that approaches that combine quantitative and qualitative methods together are more fruitful when carrying out CDA (Baker *et al.* 2008). To this end, we also decided to perform more detailed analyses based around two frequent linguistic contexts that contained the word *Muslim*. An examination of two word clusters that contained *Muslim* found that the two most frequent were *Muslim world* and *Muslim community*. Both terms refer to Muslims as belonging to a group, and it was decided to conduct further analyses of these terms by examining concordance lines and attempting to uncover common patterns or prosodies associated with them. We were particularly interested in seeing whether the terms referred to Muslims as a homogeneous set of people or whether journalists attempted to reference diversity (e.g. by using phrases like *differences within the Muslim community*). This stage of our analysis became more qualitative in that we examined full articles rather than concordance lines, and at times we took an intertextual approach by considering how articles sometimes referred back to earlier articles (see the discussion of piggy-banks below).

Noun collocates of Muslim

Using Sketch Engine, the word sketch of *Muslim* revealed that, as it tended to be used in its adjectival form in the corpus (70% of the time), it collocated most frequently and saliently with nouns. *Muslim* modifies 1,256 different

nouns (types) in 84,671 instances (tokens).[9] A first interesting observation is that a mere 2% of these noun types (the 25 most frequent collocates) account for >50% of all noun tokens modified by *Muslim* (Table 1).

This small, but highly frequent, subset of noun collocates provides strong initial indications regarding the main topics indexed by the use of *Muslim* as an adjective, such as conflict and violence (*extremist, fanatic, soldier, terrorist*), the view of Islam in terms of ethnic or national identity, or, more generally, as an homogeneous organized entity (*community, world, country, leader, state, nation*), with its own unique socio-cultural attributes (*woman, man, girl, family, youth,*

Table 1: Most frequent noun collocates of adjective Muslim

Noun collocate	Frequency	Salience
community	7676	10.4
world	4928	9.19
woman	4484	9.07
country	3864	8.74
leader	3217	8.8
cleric	2279	9.36
man	2158	7.5
group	1837	7.7
population	1799	8.9
extremist	1710	8.8
school	1220	7.7
state	997	7.4
girl	966	7.8
family	901	7.1
faith	823	7.7
nation	805	7.7
organization	696	7.5
youth	630	7.6
fanatic	612	7.7
student	579	7.2
soldier	519	6.5
child	505	6.2
people	489	4.9
terrorist	467	6.7
pupil	452	7.2

Salience is a measure of the strength of collocation and is calculated by the *logdice* statistic, see Curran (2004).

student, child, pupil), and, to a lesser extent, religion (*cleric, faith*). However, it is possible that if less frequent collocates are also considered, a different picture may be revealed (Baker 2004). Therefore, all 1,256 collocates were examined and put into thematic categories through manual concordance analysis. This revealed that the noun collocates of the adjective *Muslim* belong to a small number of categories (summarized in Table 2).

It must be clarified that the categorization does not rely so much on the dictionary meaning of the noun collocates, as on the topics they index in the corpus articles. For example, the word *shop* may not immediately appear to fit into the category of 'characterizing/differentiating attributes', but when we consider how it is regularly used in the context of 'Muslim shop', it appears to have a differentiating function.

What emerges from the examination of the categories in this section is the presentation of Muslims as a homogeneous population embroiled in conflict, either as aggressors or victims—most frequently the former. In turn, issues of conflict are presented (explicitly or implicitly) as emanating from their religion (Islam). The categories are discussed in more detail below.

Table 2: Categorized noun collocates of adjective Muslim

Categories and sub-categories (when applicable)	Examples of noun collocates
Conflict	*extremist, fanatic, terrorist, fundamentalist*
Religion	*cleric, faith, festival, preacher*
Culture	
Social practices	*dress, culture, name, tradition*
Education	*school, teaching, education, college*
View/attitude/emotion	*opinion, anger, voice, attitude, grievance*
Ethnic/national entity	
Population	*community, population, nation, world*
Area/country	*country, state, area, region, land*
Governance	*leader, voter, MP, government, ruler*
Characterizing/differentiating attributes	
Age/sex	*woman, man, girl, youth, child, teenager*
Family/relationship	*family, parent, brother, friend, wife*
Occupation/role	*officer, patient, doctor, worker, assistant*
Ethnicity/race/nationality	*Briton, Albanian, Malay, Arab*
Other	*house, shop*
Group/organization	*group, organization, association, charity*

Conflict.   All nouns in this category relate to conflict, irrespective of whether it manifests itself in the form of verbal argument (ranging from problem-solving-oriented discussion, to verbal confrontation, to inflammatory comments, to incitement to violence) or violent/armed conflict and the related issues of damage and death. Types belonging to this category have predominantly negative meaning (e.g. *extremist*); however, related positive words are also included, as they are used in the discussion of 'negative' topics. For example, the collocation *Muslim moderate(s)* is used within discussions of fundamentalism or extremism to juxtapose the two stances (explicitly or implicitly). Simply put, discussion of moderation or peace arises when conflict or extremism becomes an issue. The excerpt below exemplifies this category, whereas it also provides a first indication of how such references tend to cluster and reinforce one another.

> To make matters worse, the head of the Muslim Council, Iqbal Sacranie, has been screaming about how we keep referring to the terrorists who killed 54 people and injured 700 more as **Muslim extremists**: ''Why not just call them criminals?'' he demanded. And that's precisely the kind of idiocy that gives root to the political correctness that has allowed this country to become the world headquarters for Islamic terrorism. For Mr Sacranie's benefit, let me explain that the reason we call these murdering bastards **Muslim extremists** is because that's precisely what they are. They're not decent **Muslim moderates**, they're extremists who kill and maim people here and around the world in the name of Islam. (*Sunday Mirror*, 17 July 2005)

Religion.   The collocates in this category refer to religious beliefs and practices relating to Islam. It must be made clear, however, that they do not normally appear in neutral contexts, that is, in balanced discussions of religious matters or informative articles on the religion of Islam, but are normally caged in discussions related, at least indirectly, to conflict:

> Hizbollah's move was denounced yesterday by Mr Siniora as a ''coup'', and the prime minister went on television to declare that he refused to back down. But, faced with the prospect of fighting a bloody civil war against a militia whose Shia **Muslim faith** is shared by many of its soldiers, the army announced it would restore the sacked security chief and merely look at Hizbollah's telecommunications system. (*Sunday Telegraph*, 11 May 2008)

Culture.   This category comprises nouns relating to culture seen as 'a fuzzy set of attitudes, beliefs, behavioural conventions, and basic assumptions and values that are shared by a group of people, and that influence each member's behaviour and each member's interpretations of the ''meaning'' of other people's behaviour' (Spencer-Oatey 2000: 4). Again, as the example below exemplifies, references to culture are not isolated, but are interweaved with issues of

conflict—irrespective of the attitude towards Muslims expressed or reported in the article.

> But Mr Hockman told the *Daily Express*: ''Given our substantial Muslim population, it is vital that we look at ways to integrate **Muslim culture** into our traditions. Otherwise we will find that there is a significant section of our society which is increasingly alienated, with very dangerous results.'' (*The Express*, 26 November 2008)

Ethnic/national entity.   The collocates in this category index the presentation of Muslims as a collective entity in terms of ethnicity and nationality, and the related feature of governance. What is notable in this category is that the religious aspects of being a Muslim tend to be subsumed within, or treated as interchangeable with, a corresponding ethnic or national identity. For example, below, *Somalian community* (*sic*) and *Muslim community* are used interchangeably.[10]

> The death in the early hours of Sunday 12 November has caused outrage among leaders of the 1,000-strong Somalian community in Southampton. In the five years since Somalis started settling in the south coast port, the **Muslim community** has kept a low profile, but a statement by the trustees of the Southampton Somali Community Association expressed their obvious anger: ''The Somali community of Southampton are outraged at what they see as the latest of a long line of injustices against them.'' (*The Independent*, 4 December 2006)

Characterizing/differentiating attributes.   This category is intimately connected to the previous two, as the sketch collocates index an approach to Muslims, which appears to take for granted the attributes indexed in them, and which are typically used to differentiate 'Muslim' attributes from 'British' or 'Western' ones. That is, collocations, such as *Muslim doctor* or *Muslim family*, are treated as shortcuts for ethnic, national, or, more frequently, cultural attributes that can be expected of the entities referred to—clearly exemplified in the next excerpt.

> Two men have appeared in court after the suspected lover of a married Muslim woman had acid thrown in his face and was stabbed twice in the back in a possible ''honour attack''. [...] The two accused, aged 19 and 25, come from East London and are charged with attempted murder. [...] Both men are said to be related to the woman. One is understood to be her brother. [...] A neighbour said the family were ''an ordinary **Muslim family**''. (*The Times*, 24 July 2009)

Group/organization.   This category is similar to the previous one, in that it characterizes entities, albeit collective ones. However, the collocates merit a separate category because they index collective aims and their active pursuit,

rather than typical attributes. In that respect, this category is connected to that of 'ethnic/national entity'. When reference is to actual organizations, the collocation itself is 'neutral' (i.e. it simply indexes the particular organization). However, the surrounding discourse is frequently one of conflict:

> The council said it thought that Muslim voters could be decisive in deciding the result in as many as 40 to 50 seats. [...] The **Muslim Association** of Britain has said that any MP who voted against the war and against the recent anti-terrorist legislation deserved support. Any MP who backed the war and backed the anti-terrorism Bill should be opposed, it said. (*The Daily Telegraph*, 20 April 2005)

The (qualitative) concordance analysis of the sketch of *Muslim* as an adjective is complemented by the quantitative analysis of the relative frequency of the categories of use. Figure 2 shows the relative collective frequency of sketch collocates in each category, which points at the relative frequency of topics in the corpus articles. What becomes immediately apparent is that, collectively, sketch collocates indexing references to religious aspects are a minority (8.7%). The two most frequent categories are those presenting the adjective *Muslim* as an attribute of ethnicity or nationality (37.6%), or, more generally, a shortcut to typical, and differentiating, attributes (21.7%). Given the affinity of the category of 'culture' (12%) to these two categories, it then becomes apparent that almost three-quarters (71.3%) of the collocates of the adjective *Muslim* in the corpus articles seem to be used to present Muslims as a homogeneous group. The frequency of the 'conflict' category may seem fairly low (14.4%); however, its frequency becomes significant if we consider that the search query used to elicit articles consisted of terms which are, at least nominally, descriptive of the religion, its adherents, and attendant practices (e.g. related to worship). In other words, the search query was not loaded to elicit articles about conflict. Such articles appeared because that is the way the British press writes about Islam. Additionally, as shown in the excerpts previously, discourses of conflict are pervasive in the corpus articles, and are indexed by, and interweaved with, all other categories.

However, we also need to consider the possibility that the picture emerging from the quantitative breakdown of collocate tokens is the result of the extremely high/low frequency of particular types (i.e. distinct words). To this end, the frequency of types within categories and their groups was also examined (Figure 3). Additionally, the frequency distribution in terms of types provides a measure of the lexical richness of each category.

As Figure 3 shows, the frequency distribution of types reveals further details of the noun collocates of *Muslim*. The categories of 'religion' and 'characterizing/differentiating attributes' do not seem to contain many types with extremely high-token frequency, as the relative frequencies of their tokens and types are fairly comparable. However, the types in the categories of 'ethnic/national identity' and, to a lesser extent, 'group/organization' have a clearly lower relative frequency. This suggests that their respective token
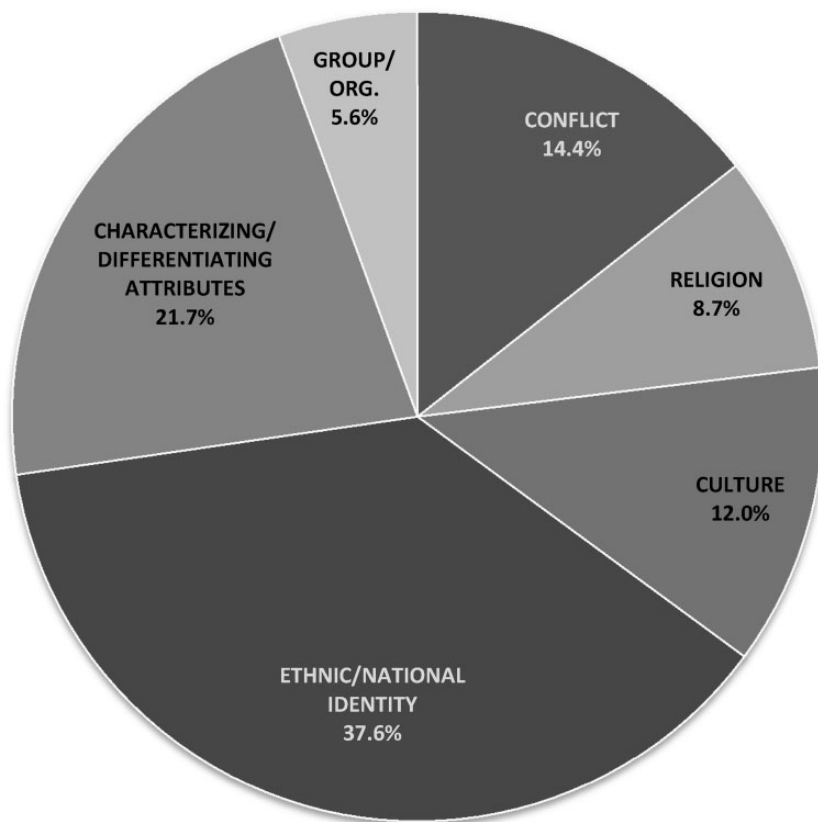
Figure 2: Frequency distribution of collocate categories in terms of tokens

frequencies depend on the repeated use of a relatively small number of types (see Table 2 for examples). Conversely, 'conflict' and, to a smaller degree, 'culture' have higher type than token frequencies, with type relative frequencies being 80 and 50% higher than token ones. This indicates that the latter two categories are particularly rich lexically, which, in turn, suggests that these topics are salient in the corpus articles.

Further patterns can be revealed by comparing the categories in terms of frequency and lexical richness. In Figure 4, the relative frequencies of types and tokens in each category have been plotted, with the intersection of the dotted lines representing the average relative frequency. Categories depicted above/below the horizontal line have relative frequency above/below average respectively, whereas those depicted on the left/right-hand side of the vertical dotted line have low/high lexical richness, respectively. Together, the dotted lines divide the plot into four quadrants representing four different combinations of frequency and lexical richness. The only category showing high
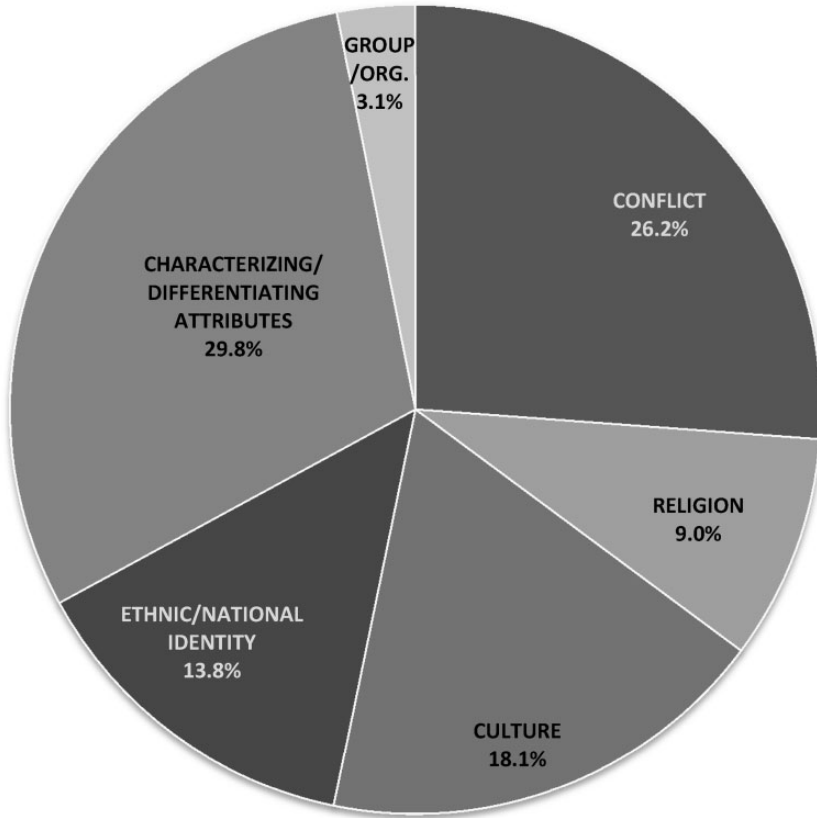
Figure 3: Frequency distribution of collocate categories in terms of types

frequency and lexical richness is that of characterizing/differentiating attri-
butes, whereas the category of religion is infrequent and lexically poor (both
values are clearly below average). The category of conflict is lexically rich,
whereas it has a relative frequency close to the average. Conversely, the cat-
egory of ethnic/national identity is extremely frequent, whereas it has lexical
richness close to the average. Combined, the aforementioned observations
strengthen the conclusion that, through the use of the adjective *Muslim* in
the corpus articles, Muslims are presented as a uniform population, typified
by its unique customs. Irrespective of its accuracy, this presentation is neither
positive nor negative. However, the presentation is rendered negative because:
(i) the putative attributes ascribed to Muslims are contrasted (explicitly or
implicitly) with those of non-Muslim UK citizens, and the West in general,
and, more importantly, (ii) this contrast is caged in the presentation of the
putative Muslim identity as intimately related to conflict.[11]
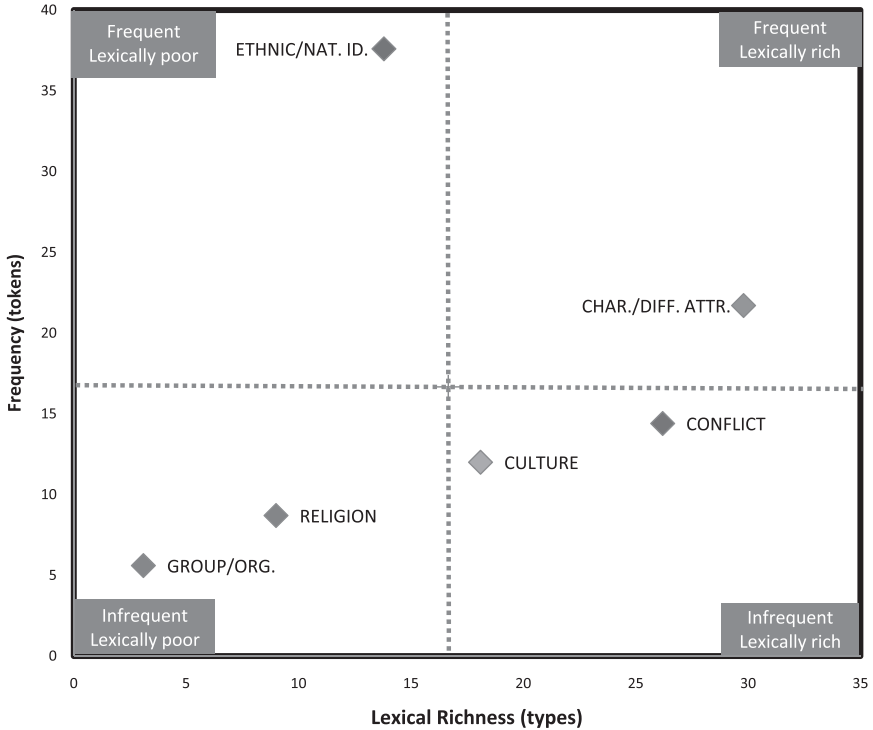
Figure 4: Comparison of frequency and lexical richness between categories

Muslim world and Muslim community

Although the aforementioned analysis has given a view of the general trends, we feel that the research would be complemented by an approach that examines a smaller number of frequent collocates of *Muslim* in more detail. Therefore, we move on now to examine the two most frequent immediate right-hand noun collocates of *Muslim*: *community* and *world*. Figure 5 shows changes in frequency over time of *Muslim community* and *Muslim world* (instances of the plural form *Muslim communities* were not included in this figure, as our focus was the frequency development of the two ways that Muslims were presented as a coherent group).

Perhaps unsurprisingly, concordance analyses found that directly after the '9/11' attacks on USA in 2001, the British press were concerned with the notion of the *Muslim world* (e.g. Muslims worldwide), whereas after the '7/7' attacks on the London transport network in 2005, there was more focus on *Muslim community*, which we hypothesized was used to refer to British Muslims.

To test this hypothesis, we categorized 100 randomly selected[12] concordance lines of *Muslim community* depending on whom they referred to. Seventeen
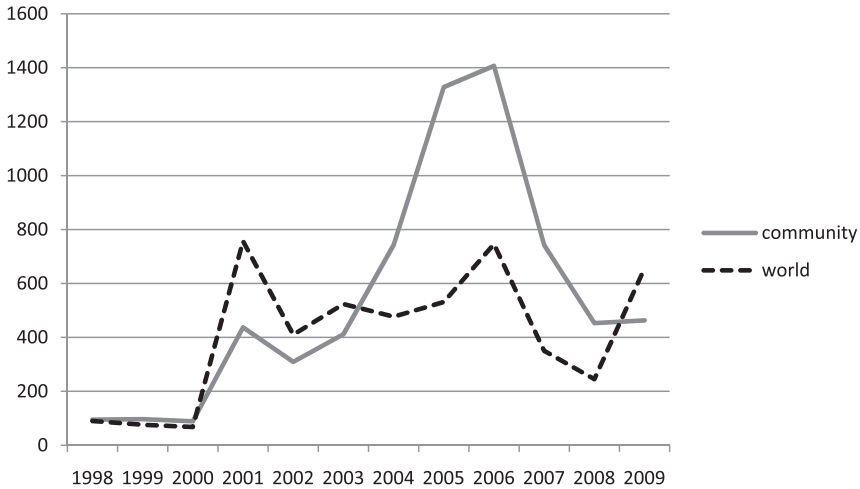
Figure 5: Change in frequency over time of Muslim community and Muslim world

cases referred to all of the Muslims in a particular town or city, for example, *the Muslim community in Bradford*. Seventy-eight referred to all Muslims living in the UK, for example, *the British Muslim community*, whereas five referenced all Muslims globally, for example, *the international Muslim community*. If these figures are representative, it then seems that *Muslim community* is mostly used in the UK press to refer to British Muslims generally.

   In the main, *Muslim community* and *Muslim world* were used uncritically to group large numbers of Muslims together. Both terms tend to be preceded by the definitive article *the* (67% of the time for *Muslim community* and 85% of the time for *Muslim world*). Examples that problematize the term, like the one below, are extremely rare:

> . . . the idea of a one-size-fits-all ''Muslim community'', or a commonality of experience, is a myth . . . Some talk about ''the Muslim community'', others feel there is no such thing. (*The Times*, 21 March 2009)

   There are only 26 cases of *Muslim community* in the corpus occurring within scare quotes, most of which express similar scepticism about the term, to note the plurality of Muslim experience. However, the British television presenter Jeremy Clarkson, writing in *The Sun*, uses the term in scare quotes in a different way during a column criticizing Tony Blair's impact on Britain. Clarkson complains that:

> In fact, we can no longer upset any ''community'' which is why the ''Muslim community'' was allowed to parade through London

urging passers-by to blow up a skyscraper and behead the infidels.
(*The Sun*, 30 June 2007)

Clarkson's use of '*Muslim community*' is ambiguous. Does he imply that such Muslims claim to belong to a community but do not really represent them, or is he suggesting that the term is used to give legitimacy to people who incite terrorism? The association of a monolithic term like 'Muslim community' (even in scare quotes) with terrorism and beheading is worrying, although Clarkson's position seems to be at the extreme end of what is considered printable for British journalism.

To find cases where Muslim communities were described as more diverse, we conducted a concordance search of *Muslim communitie*s (1,610 cases). However, this construction did not normally indicate diversity. An examination of 100 randomly selected concordance lines found that 95 were used in constructions like 'fighting between Christian and Muslim communities' or to collectivize multiple Muslim communities across the UK or the world, for example, 'Muslim communities across Britain'. Only five actually referred to Muslim communities as different from each other, whereas one of these was used to make an explicit case for 'replacing ''the Muslim community'' with ''Muslim communities''' as 'sensible and ethically sound, emphasizing the plethora of diversities' (*The Guardian*, 15 July 2008).

We then conducted searches of *within the Muslim community* (244 cases) and *section(s)/element(s)/part(s)/member(s)/segment(s)/sector(s)/ of the Muslim community* (353 cases). Although these rarer clusters did tend to refer to heterogeneity, it was generally to construct certain members of the community as dangerous:

> I am also sometimes confronted by those who point out that there are elements within the Muslim community who pose a threat to our very security. (*The Express*, 26 November 2008)

> It is undeniable that certain sections of the Muslim community are hotbeds of fundamentalism and misogyny, perhaps to a greater extent than any other modern religion. (*The Observer*, 25 June 2006)

Further examination of concordance lines reveals two clear discourse prosodies surrounding *Muslim community*. The first occurs with collocates[13] like *anger* (35 cases), *fear* (33), *concern* (25), *warned* (20), *alienating* (20), *upset* (19), *alienated* (16), *criticism* (15), *alienation* (13), *unrest* (12), *outrage* (11), *alienate* (11), *offensive* (11), *offence* (11), *uproar* (7), and *antagonizing* (6). These collocates construct the Muslim community as having the potential to be offended. Concordance 1 shows a small sample of such cases.

Although there are many cases that portray the Muslim community as angry or offended, a subset of these stories is interesting, in that they describe other people, typically non-Muslims, as imposing bans on items because of oversensitivity. Such stories, about 'politically correct' officials or other people banning certain things on behalf of Muslims who are then described as *not* offended,

we're so afraid of antagonizing the	Muslim community	that we turn our anger on our
They're scared of anything that might go down badly with the	Muslim community	.
might be offensive to some in the	Muslim community	and that it could ''incite acts of violence
The broadcaster also enraged sections of the	Muslim community	and the West Midlands Police
move would cause resentment in the	Muslim community	.
in a dawn raid that caused uproar in	Muslim community	.
to avoid upsetting Muslims, the	Muslim community	will find it insulting because

Concordance 1: Muslim community as potentially offended

tend to occur in the right-leaning popular press:

> PIGGYBANKS are facing the axe—because some Muslims could take offence. Britain's top High Street banks have ruled the money-boxes are politically incorrect. But last night the move sparked snoutrage. And one of Britain's four Muslim MPs, Khalid Mahmoud, said: ''A piggybank is just an ornament. Muslims would never be seriously offended.'' (*The Star*, 24 October 2005)

Interestingly, despite the fact that this article stresses that Muslims would not be offended, such stories were interpreted by readers somewhat differently. The following day (25 October 2005), *The Star*'s 'Text Maniacs' column (which prints text messages from members of the public) contained several messages about the proposed banning of piggybanks:

> Y shud we change r way of life just 2 stop offending muslims. they aint neva gonna change theirs. Maybe they shud try eating pork. a nice bacon sarnie cud change any1's mind.

> muslims r offended by our piggy banks!? Then the £56 me n ma wife n ma 4 girls have got in our piggy bank 2 help the ppl in pakistan wil b spent on a fry up.

This misinterpretation of the original story casts Muslims as easily offended and paints them, rather than the alleged bank bosses, as being oversensitive. Although *The Star* also included some text messages from Muslims who were not offended by piggybanks, it could be argued that the publication of the aforementioned messages would mislead readers who had not carefully read the previous day's article, as the text messages give the impression that Muslims are offended by piggybanks. However, with such a large number of references in the press to outraged responses within 'the Muslim community', it is perhaps unsurprising that some readers misinterpret such stories.

A second discourse prosody of *Muslim community* concerns the view of the Muslim community as separate from the rest of Britain. This is most frequently referred to by the lemma ALIENATE[14] (57 occurrences), but also includes less

Britain's record in integrating the	Muslim community	and in fostering a secure, strong
creating a wedge between the	Muslim community	and mainstream society.''
to assist the assimilation of the	Muslim community	.
by the simmering conflict between the	Muslim community	and the white British.
is too little understanding between the	Muslim community	and the wider community, and vi
to help officers interact better with the	Muslim community	across Sheffield.
could kickstart its engagement with the	Muslim community	.

Concordance 2: Muslim community as separate

frequent terms like *non-assimilation*, *driving a wedge*, *too little understanding*, and *conflict*. There are also references to actions that need to be taken or are being taken to encourage Muslims to INTEGRATE (53) and ENGAGE (39) and assimilation, interaction, and understanding between the Muslim community and 'the wider community' (see Concordance 2).

Other common topics associated with Muslim communities are the notion of a backlash or attacks on them, concerns about relations between such communities and non-Muslims, and the extent to which such communities are alienated, resulting in extremism, radicalization, and terrorism:

> ''The war on terror has had a devastating effect,'' he said. ''We have become targets of the security apparatus and are seen as an enemy within unjustifiably. This has resulted in a backlash against the Muslim community. We have become the hidden victims.'' (*The Guardian*, 23 April 2005)

Bearing in mind these representations of *Muslim community*, we now move on to the term *Muslim world*. This term (occurring 4,928 times) is used with less variation than *Muslim community* in the corpus; it tends to simply refer to all Muslims across the world, particularly countries that have significant populations of Muslims.

> ''We would be doing it regardless of religion,'' he said. ''But it does give the Muslim world and the rest of the world an opportunity to see America generosity, American values in action.'' (*Daily Telegraph*, 6 January 2005)

> Iran, Iraq and Syria have been the only Muslim countries to condemn the US and British bombings, though the mood on the streets in much of the Muslim world has been hostile. (*The Guardian*, 11 October 2001)

One way to understand more fully how *Muslim world* is used, is to see what other concepts it occurs with. Examining the phrase *Muslim world and*,

emblematic of the rift between the	Muslim world and the West	, a cause that is adopted
damage'' to relations between the	Muslim world and the west	, Mr Sacranie added.
bridges between a bitterly resentful	Muslim world and the West	; between an anti-Americ
such heightened unease between the	Muslim world and the West	, and when many Turks
city's bid to build bridges between the	Muslim world and the West	. He said 'A unique
help to heal the divisions between the	Muslim world and the West	. He asked 'Is it really goin
increase in tension between the	Muslim world and the West	. There would also be the
deepened the alienation between the	Muslim world and the West	– with consequences not
. It will define relations between the	Muslim world and the West	.'' Turning to the
better understanding between the	Muslim world and the West	. In Cambridge, the HRH

Concordance 3: *Muslim world and the West*

references to the West occurred 63 times (a random sample of 10 of these cases is shown in Concordance 3).

As with the British Muslim community's relationship to the 'wider community' in Britain, the Muslim world here is characterized in terms of poor relations, with terms, such as *rift*, *bitterly resentful*, *heightened unease*, *divisions*, and *increase in tension*, directly referencing this, and phrases, like *build bridges*, *better understanding*, and *help to heal*, which imply that there are poor relations that need to be overcome. It is also worth noting a disparity in the labelling of Muslims as belonging to a 'world' and 'the West', which is simply named as a point on the compass (both are culturally relativistic terms). The term *Muslim world* implies something different from the default world of 'the West', which does not seem to require being labelled as a 'world'. Although *Muslim world* occurs 4,928 times in the corpus, in fact, the term *western world* occurs 1,051 times; therefore, it is rarer, although still reasonably frequent. However, there are 25,000 references in the corpus to *The West*,[15] so *western world* is clearly a minority choice of wording.

A question arises regarding whether other identity groups are labelled as belonging to a separate and specific 'world' in general English. To answer this question, it is useful to step outside our specialized corpus and refer to a different corpus, composed of a wider range of texts. The ukWAC[16] corpus consists of ~1.5 billion words of general English (collected through internet sources) and is also available through the Sketch Engine interface. Common left-hand descriptors of *world* in this corpus were *Arab*, *Western*, *English-speaking*, *Islamic*, and *Muslim*. Other religions, such as *Jewish*, *Christian*, *Catholic*, and so forth, were not commonly used with reference to *world*. Also, terms that distinguish between different branches of Islam (such as *Sunni* or *Shia*) were not common as collocates of *world*. Instead, *Muslim world* seems to construct Muslims as belonging to a homogeneous group that is distinctly separate from non-Muslims. Clearly, the same process is at work with terms like *English-speaking world* or *Western world*, but rather than defining

people by religion, members of these 'worlds' are instead defined by the language that they speak or where they live (even though many Muslims actually reside in 'The West').

The term *Muslim world* has been criticized by Carpenter and Cagaptay (2009), who write that it 'is not only an analytical error—it's also a critical public diplomacy mistake... Muslim world unfairly and singularly assigns adherents of Islam into a figurative ghetto. And particularly... post-September 11, this relegation carries a real moral hazard. Extremists are the only Muslim group that strongly advocates tying all Muslims together politically, in a united global community... Every time the United States speaks to the Muslim world, then it inadvertently legitimizes the extremists' vision.'

It is questionable whether journalists and others who use *Muslim world* are aware of this critical reading of the term (one of the authors of this article previously used it uncritically in our pilot study, Baker [2010: 328]). Additionally, we should note that there are other potential interpretations of the term, and that it is incorporated in the name of the Muslim World League, one of the largest Islamic non-governmental organizations, whereas there are journals entitled *The Muslim World* and *Muslim World Journal of Human Rights*. Perhaps then, the acceptability of the term depends on the context of its use. For example, if it is used repeatedly to suggest that there are problems between 'the Muslim world' and 'the West' then the sense of two disparate entities will be exaggerated. As suggested by Concordance 3, such a construction does seem to be frequent in the corpus. Frequent collocates of *Muslim world* include references to the 'other' (*west* [149], *western* [83], *America* [58], *American* [47], *British* [23], *Britain* [38]) relationships (*relations* [70], *relationship* [22], *between* [132], *against* [117]), and words which indicate conflict (*war* [49], *protests* [36], *anger* [36], *outrage* [29], *attack* [27], *fury* [20], *hatred* [20], *conflict* [22]).

CONCLUSION

Our approach has focused on a large-scale analysis of the word *Muslim* in the British press. Although the corpus approach helps to reduce the potential for researcher bias to influence our analysis, it is worth reflecting on some limitations of the study.

First, we cannot claim that we have approached the analysis from a completely unbiased position. We would argue that anybody who makes such a claim is misguided—all social research is biased (see Burr 1995). Instead, we acknowledge that, to an extent, our interpretations of data are still led by aspects of our own identities (such as being non-Muslims) and cognitive biases that humans generally possess. We have tried to adopt a reflexive mindset and share our findings with a range of other people from different backgrounds, and this has helped us when framing our interpretations (e.g. we initially viewed *Muslim world* as always problematic, until conversations with Muslim colleagues suggested that it had a range of meanings).

Secondly, we can make no claims about the ways that readers internalize the repeated sets of associations (the prosodies) in newspapers. Readers are not passive 'dustbins', and McIlvenny (1996) argues that meaning is created from interaction between a text and its readers. Hall's (1973) notion of 'resistant readers' is also relevant, particular in an age where most newspaper articles are now freely available and circulated online. We can no longer view newspapers as only having a loyal readership who make the decision to regularly purchase that periodical. Instead, controversial stories can be widely circulated online through sites like Twitter and Facebook, meaning that the concept of 'audience' encapsulates almost everyone, including many people who would never choose to read that newspaper, but may be galvanized to complain about certain articles.

Overall, the quantitative analysis found that Muslims were frequently constructed in terms of homogeneity and connected to conflict. This finding was echoed in the qualitative analysis that showed that *Muslim community* and *Muslim world* were frequently characterized as distinct, reasonably homogeneous entities that are quick to take offence, in tension with the UK or 'the West', rather than integrated, contain dangerous radical 'elements', and are threatened by a backlash. The terms help to create the idea of Muslims as belonging to a distinct and separate 'imagined community' at both the global and national level, and, thus, contribute towards a process of 'othering'. Even articles that could be viewed as 'well meaning', in that they stress the importance of integrating the Muslim community, rather than demonizing Muslims, still contribute towards this othering process.

The qualitative analysis also shows how some newspapers print letters (or text messages) from readers that contain negative representations of Muslims. The publication of such text messages could be viewed as a legitimation strategy, in that the newspaper can distance itself from such opinions, whereas at the same time it can claim to give its readers a voice. As contact details of these letters and messages are not printed, it is difficult to make such readers accountable, although we would argue that ultimately the newspapers themselves should take responsibility as they have editorial control. The text messages regarding the banning of piggybanks show how a misinterpretation of an original news story was legitimized by the paper that printed the original story, resulting in Muslims being cast in a more negative light the following day. Not only did such text messages present an inaccurate and misleading view of the original story; statements in them, such as that Muslims should try eating pork, could be interpreted as strongly provocative. In this case, we would question the responsibility of the newspaper that printed these texts.

A further legitimation strategy involves the use of columnists, such as Jeremy Clarkson, whose more extremely negative views can also be distanced from the newspaper's overall stance. The British press is self-regulated by the Press Complaints Commission, and in the past, it has rejected complaints about Islamophobic columnists (see Petley 2006: 56), noting that such columns are only representative of the columnist's own opinions. Again, we would question whether the distinction between columnist and newspaper is so clear-cut.

Newspapers choose to publish what the columnists write and such columns may, thus, represent a 'common sense' position to some readers. The capacity for negative representation is, thus, enshrined within the newspaper's discursive structure and the discursive systems that British journalism operates under. We, thus, cautiously welcome the closure of the Press Complaints Commission in 2012, and hope that its replacement will take a different approach to complaints about Islamophobic 'opinion'. We would particularly advocate that changes to the Editor's Code of Practice are implemented to cover such cases.

Finally, it will be interesting to observe whether the internet and social networking, with their potential to alert many more critical readers to controversial stories and to afford the opportunity to complain, will impact on the way that newspapers represent Muslims in the coming years.

Conflict of interest statement. None declared.

NOTES

1   In the UK, two types of newspapers are distinguished as follows: *broadsheets* and *tabloids*. Broadsheets are 'considered to contain serious, in-depth journalism', whereas tabloids are 'usually characterized as popular in style and dominated by sensational stories' (definitions from the *Oxford English Dictionary*, online edition: http://www.oed.com).

2   From the Office for National Statistics http://www.statistics.gov.uk/cci/nug get.asp?id=954.

3   Grant number RES-000-22-3536.

4   http://www.guardian.co.uk/media/ table/2009/jan/09/abc-december-national-newspapers.

5   For example, see the sixth report of the committee (2004/2005) and the sixth report of the committee (2007/ 2008).

6   The authors would like to thank the Nuffield Foundation for the grant that supported this work.

7   Wodak and Meyer (2009: 19-20, 22) include the 'Corpus Linguistics Approach' as a research strategy recognized by CDA. They stress that 'CDA has never been and has never attempted to be or to provide one single or specific theory. Neither is one specific methodology characteristic of research in CDA' (ibid: 5). Although work in CDA can benefit from the use of large corpora (see e.g. the 2011 special issue of *Critical Discourse Studies* edited by Michal Krzyzanowski), it could be argued that not all genres lend themselves equally to corpus analysis, and additionally, some CDA practitioners would only view CL as a method for supporting qualitative discourse analysis, such as downsizing large corpora of media texts, rather than as a full approach to CDA.

8   These newspapers were *The Sun, The News of The World, The Mirror, The Sunday Mirror, The Star, The Sunday Star, The Daily Express, The Sunday Express, The Mail, The Mail on Sunday, The People, The Guardian, The Observer, The Independent, The Independent on Sunday, The Daily Telegraph, The Sunday Telegraph, The Times*, and *The Sunday Times*. One weekly magazine, *The Business*, was also included.

9   Word sketch combines singular and plural instances of the same word together, so in Table 1, the term

community refers to *community* and *communities*.

10 This treatment is similar to the conflation in use of the terms *refugees*, *asylum seekers*, and *immigrants* reported in Baker *et al.* (2008) and Gabrielatos and Baker (2008).

11 One aspect of representation we also considered was to do with differences in the presentation of Muslims between text written by journalists compared with quoted speech from other people. We examined a random sample of 100 concordance lines of *Muslim*. Only 12 were within quotations and, in those instances, Muslims were not constructed negatively. Clearly, further research is needed on the matter.

12 All random samples were generated by Sketch Engine's 'random sample' function.

13 Collocates in this section were calculated as raw frequencies within a range of five words either side of the search word—as is common practice in corpus linguistics (see also Sinclair 1991: 105)—occurring five times or more.

14 Terms in small capitals refer to all grammatically related forms. Thus ALIENATE denotes the words *alienate*, *alienated*, *alienates*, *alienating* and *alienation*.

15 This excludes cases like *the West Bank/ Midlands/Indies/coast/end/country*.

16 UKWaC stands for *UK Web Archiving Consortium* (see Bailey and Thomson 2006).

REFERENCES

Akbarzadeh, S. and **B. Smith**. 2005. *The Representation of Islam and Muslims in the Media (The Age and Herald Newspapers)*. Monash University.

Awass, O. 1996. 'The representation of Islam in the American media,' *Hamdard Islamicus* 19/3: 87–102.

Bailey, S. and **D. Thompson.** 2006. 'UKWAC: building the UK's first public web archive. D-Lib Magazine, 12/1', available at http://www.dlib.org/dlib/january06/thompson/01thompson.html. Accessed 25 September 2012.

Baker, P. 2004. 'Querying keywords: questions of difference, frequency and sense in keywords analysis,' *Journal of English Linguistics* 32/4: 346–59.

Baker, P. 2010. 'Representations of Islam in British broadsheet and tabloid newspapers 1999-2005,' *Language and Politics* 9/2: 310–38.

Baker, P., C. Gabrielatos, M. Khosravinik, M. Krzyzanowski, T. McEnery, and **R. Wodak.** 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 and Society* 19/3: 273–306.

Bloor, M. and **T. Bloor.** 2007. *The Practice of Critical Discourse Analysis: An Introduction.* Hodder Arnold.

Burr, V. 1995. *An Introduction to Social Constructionism.* Routledge.

Carpenter, S. and **S. Cagaptay.** 2009. 'What Muslim world? Foreign Policy', available at http://www.foreignpolicy.com/articles/2009/06/01/what_muslim_world. Accessed 2 June 2009.

Conboy, M. 2010. *The Language of Newspapers: Socio-historical Perspectives.* Continuum.

Curran, J. R. 2004. 'From distributional to semantic similarity,' Ph.D. thesis, Edinburgh University.

Dunn, K. 2001. 'Representations of Islam in the politics of mosque development in Sydney,' *Tijdschrit voor Economische en Sociale Geografie* 92/3: 291–308.

Fairclough, N. 1995. *Critical Discourse Analysis.* Longman.

Gabrielatos, C. and **P. Baker.** 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.

Hall, S. 1973. *Encoding and Decoding in the Television Discourse*. Birmingham University, Centre for Cultural Studies.

Hardt-Mautner, G. 1995. 'How does one become a good European: the British press and European integration,' *Discourse and Society* 6/2: 177–205.

Kilgarriff, A., P. Rychly, P. Smrz, and D. Tugwell. 2004. 'The Sketch Engine,' *Proceedings of Euralex.* pp. 105–16.

Mautner, G. 2009. 'Corpora and critical disourse analysis' in P. Baker (ed.): *Contemporary Corpus Linguistics*. Continuum, pp. 32–46.

McIlvenny, P. 1996. 'Heckling in Hyde Park: verbal audience participation in popular public discourse,' *Language in Society* 25: 27–60.

Moore, K., P. Mason, and J. Lewis. 2008. *Images of Islam in the UK: The Representation of British Muslims in the National Print News Media 2000-2008*. Cardiff School of Journalism, Media and Cultural Studies.

Partington, A. 2006. *The Linguistics of Laughter: A Corpus Assisted Study of Laughter Talk*. Routledge.

Pearce, M. 2008. 'Investigating the collocational behaviour of *MAN* and *WOMAN* in the British National Corpus using Sketch Engine,' *Corpora* 3/1: 1–29.

Petley, J. 2006. 'Still no redress from the PCC' in E. Poole and J. E. Richardson (eds): *Muslims and the News Media*. Taurus, pp. 53–62.

Poole, E. and J. E. Richardson. 2006. *Muslims and the News Media*. I. B. Tauris.

Reisigl, M. and R. Wodak. 2001. *Discourse and Discrimination: Rhetorics of Racism and Antisemitism*. Routledge.

Richardson, J. E. 2004a. *(Mis)Representing Islam: The Racism and Rhetoric of British Broadsheet Newspapers*. John Benjamins Publishing Company.

Richardson, R. (ed.). 2004b. *Islamophobia: Issues, Challenges and Action*. Trentham Books.

Sinclair, J. 1991. *Corpus Concordance Collocation*. Oxford University Press.

Spencer-Oatey, H. 2000. 'Rapport management: a framework for analysis' in H. Spencer-Oatey (ed.): *Culturally speaking: Managing Rapport Through Talk Across Culture*. Continuum, pp. 11–46.

Stubbs, M. 1996. *Text and Corpus Analysis*. Blackwell.

Tognini-Bonelli, E. 2001. *Corpus Linguistics at Work*. John Benjamins.

Van Dijk, T. 2009. 'Critical discourse studies: a sociocognitive approach' in R. Wodak and M. Meyer (eds): *Methods of Critical Discourse Analysis.* 2nd edn. Sage, pp. 62–86.

Van Leeuwen, T. 2007. 'Legitimation in discourse and communications,' *Discourse and Communication,* 1/1: 91–112.

Widdowson, H. G. 2004. *Text, Context, Pretext. Critical Issues in Critical Discourse Analysis*. Blackwell.

Wodak, R. and M. Meyer. 2009. 'Critical discourse analysis: history agenda, theory and methodology' in R. Wodak and M. Meyer (eds): *Methods of Critical Discourse Analysis.* 2nd edn. Sage, pp. 1–33.