John Benjamins Publishing Company

A corpus-driven comparison of English and French Islamist extremist texts

Paul Baker and Rachelle Vessey

Lancaster University | Birkbeck, University of London

Using corpus linguistics and qualitative, manual discourse analysis, this paper compares English and French extremist texts to determine how messages in different languages draw upon similar and distinct discursive themes and linguistic strategies. Findings show that both corpora focus on religion and rewards (i.e. for faith) and strongly rely on othering strategies. However, the English texts are concerned with world events whereas the French texts focus on issues specific to France. Also, while the English texts use Arabic code-switching as a form of legitimation, the French texts use a formal register and quotation from scripture in discussions of permissions, rights, obligations and laws. Finally, the English texts refer to and justify violence to a greater extent than the French texts. This paper contributes to the field of terrorism studies and the field of corpus linguistics by presenting a new approach to corpus-driven studies of discourse across more than one language.

Keywords: French, English, extremism, cross-linguistic corpus analysis, corpus-driven discourse analysis

1. Introduction

Islamist extremist discourse is a difficult subject to research. The focus and handling of extremist texts are clearly dangerous (not least because possession of the texts is illegal, at least in UK law), but researchers also face the added difficulty that this form of communication is sophisticated and prolific (Frampton et al., 2017; Zelin, 2015). The terrorist groups al-Qaeda and the Islamic State of Iraq and the Levant (known as ISIL, ISIS and Daesh) view the media as an important arena for advocating jihad and employ dedicated media centres, which produce materials in multiple languages for current and potential recruits (Ingram, 2017; Novenario, 2016). Researchers have started to tackle extremist discourse using these difficult materials as data; however, the vast majority of this (still nascent) work has

focused on English language publications, despite the fact that both groups have published in multiple languages (Aggarwal, 2017; Frampton et al., 2017; O'Halloran et al., 2016; Wignell et al., 2017a, 2017b; Zelin, 2015). The relevance of comparing English and French publications seems clear when we consider the number of attacks taking place in English- and French-speaking locations (e.g. the United Kingdom, France and Belgium) (Buril, 2017). The perpetrators of attacks have either identified themselves as members of groups like al-Qaeda or ISIS, and/or such groups have claimed responsibility for the attacks. Many of the perpetrators of attacks have been identified as British and French nationals, often in possession of illegal terrorist materials; some of these individuals were radicalised as a result of communication with the groups or via the groups' media propaganda (Droogan & Peattie, 2016; Novenario, 2016). The availability of extremist materials in languages like English and French ensures that the groups have possible influence that extends far beyond the Arabic-speaking heartlands where the groups' headquarters are based geographically.

In this paper, we compare corpora of English and French Islamist extremist texts in order to determine the extent to which terrorists' messages in different languages draw upon similar and distinct themes and linguistic strategies and appear to be written for the same purposes. Our research thus contributes towards wider understandings of Islamist extremism, particularly when considering the transmission of messages across international borders and the extent to which extremist discourse is "tailored" for specific contexts. However, our paper also brings to light issues in terms of comparing corpora consisting of different languages, describing some of the methodological decisions that we made in order to achieve this. Section 2 describes a number of studies that are relevant to this research, particularly those which have used corpus techniques to examine Islamist extremism. Next, our method section describes the corpora that were used and the way that we derived and then categorised keywords. In the analysis section we outline differences and similarities between the two datasets and we then provide explanations for our findings, noting some of the ways that lexical choices could contribute towards successful transmission of extremist discourse. We conclude by offering a reflection on the methodological process we implemented and describing further research that could be carried out on the basis of our findings.

2.   The violent extremism and publications of the Islamic State and al-Qaeda

Conceptualisations of the relationship between the term 'extremism' and its relation to Islam are complex and contested. One distinction is between the terms 'Islamic' extremism and 'Islamist' extremism, where the former has been popularly used in western media (and implies a form of extremism connected to Islam generally) and the latter refers to extremism connected more specifically to politically-motivated Islam. A related term, 'Islamism' means the desire to impose a version of Islam over society while other terms like 'militant Islam', 'radical Islam' or 'fundamentalist Islam' complicate definitions further and are also often found in western news media, implying that Islam is an intolerant and insensitive religion (Esposito, 2014; Kramer, 2003; Poljarevic, 2015). Baker et al.'s (2013) research, for example, has tracked how British newspapers moved from describing Islam as hardline to fanatical to militant to radical to extreme between 1998 to 2009. Additionally, some writers have argued that such terms tend to characterise Islam as an inherently extremist religion, whereas it is people who are the problem so it is preferable to refer to *extremist Muslims* (Ghilan, 2013). The British government has defined *Islamist extremism* as any form of Islam that opposes "democracy, the rule of law, individual liberty and mutual respect and tolerance of different faiths and beliefs" (HM Government, 2013:1), while the UK High Courts have at different times linked extremist Islamic positions to the use of violence aimed at achieving perceived Islamic goals or an extremely conservative view of Islam, which does not involve violence (see e.g. [2016] EWCA Crim 568 and [2016] EWHC 2688 (QB)).

Groups like al-Qaeda and ISIS have carried out attacks on both opposing military forces and civilians, including weaponising planes and land vehicles, exploding bombs in public places, and kidnapping and beheading people, as well as engaging in more traditional forms of warfare. While there are many differences between these two groups (see e.g. Ingram, 2017; Novenario, 2016), in this paper we focus on their similar global influence, achieved at least in part through their use of online messages that aim to recruit new members and inspire them to carry out acts of violence. Indeed, an English language magazine published by al-Qaeda is called *Inspire* and, as suggested by both its title and its opening article, the purpose of the magazine is to "inspire the believers to fight" (Ingram, 2017:358). Magazines such as *Inspire* thus form part of the terrorist groups' strategy used to promote jihad internationally.

The appeal and influence of al-Qaeda's *Inspire,* which was first published in 2010, led to the creation of magazines for other extremist groups. For example, *Dabiq,* which dates from 2014, is an ISIS English-language magazine that has

been strongly influenced by *Inspire,* mimicking its use of regular sections and adopting iconography and doctrine (Frampton et al., 2017; Ingram, 2017). As the magazines have grown in notoriety and influence, the amount of research around them has increased.

Researchers have drawn on *Inspire* and *Dabiq* as data in order to examine narrative themes (Droogan & Peattie, 2016), the ISIS world view (Wignell et al., 2017b), radicalisation strategies (Ingram, 2017), and to compare and contrast ISIS, the Taliban, and al-Qaeda (Aggarwal, 2017; Ingram, 2017; Novenario, 2016). Other researchers have been interested in mapping out the evolution of the language used by ISIS more generally (Vergani & Bliuc, 2015). The methods have been equally wide-ranging, drawing on content analysis (Novenario, 2016), radical narrative analysis (Ingram, 2016), thematic network analysis (Droogan & Peattie, 2016), and multimodal social semiotics (Wignell et al., 2017a, 2017b). Although some computerised analysis has been applied to these texts (e.g. Vergani & Bliuc, 2015), to our knowledge, only a handful of examples of corpus linguistic research on Islamic extremist texts exist thus far. Conoscenti (2016), for example, examines collocates in a corpus of *Dabiq* articles in order to identify the different types of enemies mentioned in the texts. Also, Prentice and colleagues (Prentice et al., 2011, 2012) apply the online tool *Wmatrix* (which automatically assigns semantic codes to words in the corpus, see Rayson, 2009) to several different datasets of extremist texts. Their 2011 study examines 50 extremist texts which advocated violence in relation to the Gaza conflict, in order to identify persuasive strategies in two time periods. *Wmatrix* identified semantic groups of words that were collectively labelled as "Unselfish", "Belonging to a Group" and "Helping" and indicated the types of strategies used. The 2012 study compares corpora of counter-extremist and pro-extremist texts finding a number of key semantic category differences including sets of contrasting concepts like allowed/not allowed, dead/alive, calm/angry and lawful/crime in the extremist messages.

The advantages of this diverse range of research foci and methodologies notwithstanding, the work to date has focused almost exclusively on English language data (for an exception, see Aggarwal, 2017). This focus on English data creates a "biased understanding" (Zelin, 2015) and, crucially, it means that the extent to which such materials pose a threat to national security "cannot be properly understood" (Frampton et al., 2017: 23). Almost all ISIS media content is published in Arabic (Frampton et al., 2017), and although English is the second largest language of publication, it represents only a fraction of the overall jihadist output, which exists in other languages, too (see Table 1).

Much work remains to be done on the multilingual extremist texts and, we argue, corpus linguistics has much to offer as a methodology to this end. The small number of corpus linguistic studies referred to above have mainly relied on soft-

Table 1. ISIS media releases by language (adapted from Zelin, 2015)

Language	Number of media releases over 1-week sample period (April 18–24, 2015)
Arabic	123
English	8
Russian	5
Kurdish	4
French	2
Urdu	1

ware which assigns words to pre-defined categories. However, this can result in problems when the software encounters words that are not included in its lexicon or have multiple meanings. Our experiments using automated tagging on a sample of extremist literature in English found numerous cases of mis-tagging, as well as a substantial number of Arabic words that were tagged as "unmatched" (meaning the tool did not recognise them). We also struggled to find a tagger that could cope with both English and French. Given these difficulties, we found that an approach combining corpus linguistic techniques with qualitative, contextual-based analyses was best suited to the task at hand. This multi-faceted approach is outlined in the following section.

3.   Data and methods

The English corpus consists of 27 texts (679,743 words) categorised as extremist by UK authorities. Twenty of the texts are from two magazines, *Inspire* and *Dabiq*, while the other seven are individual texts from other extremist sources. The French corpus contains 9 texts (191,344 words) from the ISIS magazine *Dar al Islam*, which has also been categorised as extremist by UK authorities. While the total amount of texts might seem small in number, the 9 texts in the French corpus represent a 90% sample of the total *Dar al Islam* publication, as only 10 issues were published (Buril, 2017). Similarly, the 11 issues of *Dabiq* and the 9 issues of *Inspire* used in the English corpus represent 73% and 64% of the total *Dabiq* and *Inspire* publications, as only 15 and 14 issues were published, respectively (Droogan & Peattie, 2016, Wignell et al., 2017c). The complete volumes of these publications were not included in the corpora because they were either not available from the source (see below) or they were corrupted files that were not computer-readable.

There are obvious differences between the two corpora – not least of which their size, source texts, and the language being employed. It is also worth noting that although one corpus is English and the other is French, both contain code-

switching into Arabic, as will be discussed in later sections. There are also notable corpus-*internal* differences. *Inspire* and *Dabiq* were published by different extremist groups with different agendas and readerships (see e.g. Ingram, 2017; Novenario, 2016); also, each magazine consists of dozens of individual texts, many of which are written by different authors and contributors. Despite the inter- and intra- corpus differences, all of the texts included in the corpora were assessed as dangerous and extremist by relevant UK authorities and it is on this basis that we have included them in our corpora. These texts are based around similar topics and have similar goals (i.e. to persuade readers to join their cause); as a result, there is considerable lexical repetition meaning that millions of words are not likely required to identify linguistic patterns. Nevertheless, when dealing with relatively small amounts of data it is important to take into account cases where a word is poorly distributed and we have noted such cases in the analysis section that follows.

Although in related research some authors (e.g. Aggarwal, 2017) have sourced their extremist texts from the internet and others (e.g. Wignell et al., 2017a, 2017b) have published large-scale excerpts from the magazines, the possession of these texts can result in prosecution and their reproduction can lead to the recirculation of extremist discourse (Droogan & Peattie, 2016; Frampton et al., 2017). The legal and ethical dimensions of studying terrorist discourse are still unclear but in our case access for research purposes was made possible by an authorised government contact and in this article we provide only minimal excerpts of the data for demonstration so that the material is not recirculated.

A comparison of themes emerging from the English and French keyword lists was the main objective in our analysis. Since the resulting themes emerge from the corpora rather than from related literature, we have labelled our approach 'corpus-driven' (Tognini-Bonelli, 2001). Importantly, though, a keyword list obtained from the comparison of an English wordlist against a French wordlist would highlight linguistic (e.g. French vs. English) rather than discursive differences. The cross-linguistic nature of this project therefore required the compilation of individual keyword lists for each language dataset and the subsequent comparison of these keyword lists using both quantitative and qualitative means (see Vessey, 2013).

To compile individual English and French keyword lists, English and French reference corpora were required; however, a publicly-available corpus of general contemporary French proved impossible to find. Instead, a corpus of 2,310 texts from *La Presse* (a French Canadian newspaper) (1,462,398 words) from 2009 was used as a reference corpus. As discussed in Section 5, a corpus of European French might have been more appropriate, not only due to the focus in the French data on issues from France but also differences in register between the primary corpus and

the reference corpus; however, this was impossible to predict before undertaking the research and it speaks to the issues arising from the lack of a large general corpus of French. To ensure comparability between the two languages, we built a similar reference corpus of all British national newspaper texts from around the same time period, containing both broadsheets and tabloids and left-leaning and right-leaning newspapers (1,381,219 words).

With these reference corpora, *AntConc* (Version 3.4.4, Anthony, 2016) was used to derive keyword lists in French and English and the top 500 keywords for each list were examined using collocate, cluster and concordance analysis (established functionalities within corpus-assisted discourse studies; cf. Baker, 2006; Partington, 2004).

The collocates of each keyword were established using a search window of three spaces to the left and the right of the node word and a minimum Mutual Information score of five. As there is no agreement on an ideal span for collocation, we experimented with different settings in order to determine a span which gave the best precision and recall of true collocates (e.g. words where some sort of direct relationship with one another can be identified). Collocates helped us to identify how individual words were used in context, allowing us to establish the meaning of words and how concepts and themes denoted by words are connected in discourse. Clusters of 2–4 words to the left and the right of each keyword were also examined to investigate how meaning was cemented across multi-part units of language. Concordance line analysis was used to supplement collocate and cluster analysis.

Collocate, cluster and concordance analysis were used to identify themes emerging from the keyword lists. Themes that emerged in one language were compared against the themes that emerged in the other language and the themes were discussed by the researchers to determine if the categorisations were consistent and overlapping across the datasets. Then, the English and French keyword lists were compared in order to establish which keywords were common to both corpora and which were unique to the English and French magazines. This step required the careful consideration of keywords and their translation equivalents or near-equivalents in the other language. In order to take into account the fact that our two corpora were different sizes, we compared normalised frequencies between the two corpora, along a common baseline of occurrences per 1,000 words.

A final step involved mapping the equivalence (or non-equivalence) of keywords across languages (i.e. the cross-linguistic similarities and differences) onto the thematic categorisations previously established. This required revisiting the categories established in the previous step and also, in some cases, revisiting the qualitative analysis and revising some of the categories. The ultimate outcome of

this procedure was a matrix that cross-listed the emergent thematic categories against the keywords that were common to both languages or unique to English and French (see Table 2).

For the purposes of this paper, we will focus only on the categories that contained the most word types and/or tokens, attributing to these categories a dominant thematic role in the datasets. More specifically, we will provide detail on the dominant themes *shared* between the English and French magazines and the dominant themes that were *more salient to* the English and French magazines, respectively. Dominance was established by calculating the sum of relative frequencies of all words in these categories and the differences between these relative frequencies in English and French. Where relative frequencies were both high and similar in English and French, we consider these to be shared themes. Where relative frequencies were exceptionally higher in one language category than the other,[1] we consider these themes more salient to that language. Findings relating to dominant shared and salient categorical themes are outlined in the next section.

4.   Findings

In this section, common categories are presented first; then we present findings more salient in the English data, followed by Section 4.3, which presents findings that were more salient in the French data.

4.1   Common categories

As Table 2 shows, one of the categories with high relative frequencies in both English and French texts pertains to Allah and worship, indicating that religion is a central theme in both corpora (one often overlooked and misunderstood; Droogan & Peattie, 2016; Frampton et al., 2017).

---

1. The %diff calculation (Gabrielatos & Marchi, 2012) was used to compare the relative frequencies of words in each category across the two languages. The calculation results in a positive or negative number, with numbers closer to zero indicating similar relative frequencies. Ordering the %diffs for each category allowed us to identify categories that were most similar and different between French and English. All the categories in Tables 2 and 3 have a %diff of 50 or more (note that this number is arbitrary and should not be considered a benchmark for future studies).

The most frequent word is *Allah* (or *Allāh*)[2] which comprises 75% of English and 93% of French instances of the category. We have also included words in this category which occur within frequently-mentioned Arabic honorific phrases like *sallallahu 'alayhi wa sallam* ("blessings of God be upon Him and His family and peace"), *rahimahullāh* ("Allah have mercy upon him") and *radiyallāhu* ("Allah is pleased with him"). These phrases are often used in parentheses in the texts when referring to people who are lauded by the authors, such as Osama Bin Laden. Common clusters containing *Allah* in the English texts are *may Allah* (301 occurrences), which usually involves requests for Allah to help the extremist cause (e.g. *may Allah bless this Islamic State*) and *cause of Allah* (100 occurrences) which is used euphemistically to refer to all activities approved of by the extremist writers, including violence (e.g. *By now almost all of his companions had been killed in battle for the cause of Allah*). Other English clusters containing *Allah* also involve honorifics e.g. *All praise is due to Allah* (68 occurrences). One legitimation strategy found in the corpus is to describe Allah as supporting the extremist cause, as shown in Example (1):

(1)   But Allah knows the reality of their false claims, and our killed soldiers are in the gardens of Paradise while their dead soldiers are in the dungeons of Hellfire.[3]

Findings in the French corpus are similar in many ways. The word *préserve* tends to be used in the semi-fixed honorific phrase *Qu'Allah le préserve* ("May Allah protect him") (12 occurrences). While the most frequent cluster containing *Allah /Allâh* refers to the Prophet Muhammad (*messager d'Allah/Allâh,* 59 and 59 occurrences, respectively), one notable cluster is *la loi d'Allah/Allâh* (19 and 15 occurrences, respectively), which tends to be used to justify the rule of law that the Islamic State wishes to impose on its territory (Example (2)).

(2)   […] Evidemment, et comme pour toutes ses décisions, l'Etat Islamique se réfère uniquement à loi d'Allah
("[…] Clearly, and as with all its decisions, the Islamic State refers exclusively to the law of Allah")

---

2.   The corpora contained different spellings of words derived from Arabic. In the English corpus *Allah* occurs 4,782 times, *Allaah* twice and *Allāh* 314 times. It could be argued that the latter two terms help to establish the texts as linguistically and thus politically, religiously and culturally non-western, thereby offering subtly alternative "in-group" depictions. Diacritics and vowel doubling thus act as legitimation strategies.

3.   Examples given from the corpora are intended to be illustrative of the points made, rather than being representative of the most common use of a particular keyword. We have made the decision not to give the names of the texts as we would like the focus of this paper to be on the language use of the text producers rather than the texts themselves.

In both English and French, the frequent reference to *Allah* along with claims that the writers are acting on His behalf and know what He wants (i.e. are obeying the will or law of Allah) could therefore be seen as a way of legitimating violence and violent people who are held up as role models.

The second category which was found to be frequent in both groups and mentioned with similar proportions was "Negative Othering", which referred to people and concepts who were represented as being against the version of Islam advocated by the writers of the extremist literature. Numerous other studies on extremist literature have identified similar themes, finding the construction of a "crisis-steeped bi-polar world" (Ingram, 2017: 367; see also Novenario, 2016). "Negative Othering" was the third largest category in the English texts and the fourth largest in the French texts. Within this category, the general concept of disbelief was frequently referenced, with English texts using words like *disbelief* and *disbelievers* (82 and 254 occurrences respectively), and French texts using an equivalent set of words: *mécréance*, *mécréant*, *mécréants* (153, 117, 250 occurrences, respectively). However, particularly in the English texts, distinctions were made between different types of disbelief. *Apostates/apostats* were referred to in both languages, a concept which refers to the conscious abandonment of Islam, including conversion to another religion; the Arabic term *murtaddīn* (used in the English texts 67 times) refers to such a person while *taghut/tawāghīt* means crossing limits or rebelling, referring to someone who does not worship Allah. Another frequently mentioned concept was that of *hypocrisy/hypocrisie* and *hypocrites,* which denoted people who pretend to be Muslim but conceal their disbelief. In the English texts the Arabic term *kāfir*, *kuffar* or *kuffār* was also frequent, referring to a disbeliever or concealer of the truth (usually used to refer to non-believers, especially those in westernised countries), as in Example (3).

(3)   A year later (December 2013) I flew out again, this time a direct flight to Istanbul, Turkey; I was again stopped and questioned, however after a long three hour interview, the swine-eating filth disbelievers let me go and I was put on the next flight out (stupid Kuffãr).

The related concept *kufr* refers to non-belief. Together these four Arabic words occurred 144 times in the English texts but were both infrequent in French (e.g. *koufr*, 3; *kufr*, 2; *kâfir*, 1) and not found in the top 500 French keywords, indicating a difference at the lexical level in the way this category is realised across the two languages. Another Arabic term, *takfir* (in the English texts) involves the practice of declaring someone as a *kāfir*, while another form of disbelief was *shirk/chirk,* which is the sin of polytheism or idolatry (people who commit shirk are referred to as *mushrikīn*). The Anti-Christ is also included in the "Negative Othering" cat-

Table 2. Categories common to both English and French texts

Category	Keywords in English corpus	English		Keywords in French corpus	French	
		total frequency	relative frequency		total frequency	relative frequency
Allah and worship	alayhi, alayhis, allah, allāh, almighty, lord, radiyallāhu, rahimahullāh, sake, shahādah, shuhadā, tawhīd, walā, الله	6,729	9.89	allah, allāh, divinité, préserve, tawhīd	2,021	10.56
Negative Othering	apostate, apostasy, apostates, dajjal, deviance, deviant, disbelief, disbelieve, disbelievers, evil, falsehood, grayzone, hypocrisy, hypocrites, irjā, kāfir, kuffar, kuffār, kufr, murtad, murtaddīn, mushrikīn, others, rāfidah, rāfidī, regime, rulers, safawī, sahwah, sahwāt, sin, sins, sects, secularist, shaytān, takfir, weak, shirk, tāghūt, tawāghīt	4,765	7.01	adorateurs, apostasie, apostats, apostat, associateurs, bachar, chirk, différents, école, écoles, fornication, hypocrites, hypocrisie, idolâtres, idoles, irjā, innovations, problème, rafidites, système, laïcité, mécréance, mécréant, mécréants, péché, péchés, régime, sahawāt, tāghoūt, tawāghīt	1,665	8.39
Reward	blessed, blessing, blessings, booty, ghanimah, hereafter, Jannah, jizyah, merciful, mercy, paradise, reward, slave, slaves, wealth	1,520	2.23	bénédictions, bienfait, biens, butin, échange, esclavage, esclave, esclaves, jiziyah, miséricordieux, miséricorde, paradis, récompense, salut	450	2.36

egory, referred to as *al-Dajjal* or *Shaytān* in the English texts, who is viewed as a literal personification of everything that is against Islam, as in Example (4).

(4)   Shaytān will do anything he can to make you lazy and procrastinate in doing good deeds.

A further difference between the English and French texts in this category is the presence of the word *evil* in the English texts (220 occurrences), which had no equivalent top 500 keyword in the French texts. *Evil* is used to refer to people (e.g. *evil scholars*, *evil people*, *evil leader*) but also actions (*evil deeds*, *evil plans*) and also, as shown in Example (5), as a more abstract concept or vaguer presence.

(5)   Each town has a Shar'ah court and Islāmic police who patrol the streets enjoining the good and forbidding the evil and ensuring no one breaks the law of Allāh

In this category the French texts contain references to sex (*fornication,* 20 occurrences) and schools (*école*, *écoles*, 20 and 26 occurrences, respectively), while only a handful of references to these concepts were found in the English texts (cf. Droogan & Peattie, 2016). Fornication tends to be discussed in terms of a sin forbidden by Allah (*Allah condamne la fornication et l'interdit*). Attacks on individuals and places where this "sin" is common are justified using authoritative texts or sources (see Example (6)).

(6)   Si la fornication se répand dans une cité, Allâh donne la permission de la détruire
       ("If fornication is widespread in society, Allâh gives permission to destroy it")

According to the magazines, fornication is one of the hallmarks of a civilisation in decline, providing self-evident truth for the need to take up arms against the sinners and/or immigrate to the Caliphate. Fornication is also mentioned in passing with relation to schools, which are represented as another institution of corrupt society. The French texts argue that mixing girls and boys in schools is an open door to fornication (*la mixité entre les filles et les garçons qui est une porte ouverte vers la fornication*). Like fornication, schools are barely discussed in the English texts. In French, the majority of references to schools (*école/s*) derive from a lengthy article calling for parents to abandon the French education system, which has been secular since a 2004 law banned all religious symbols in public schools (hence references to *laïcité* ("secularism"), which tend to refer to the school system). Here, schools are referred to not only as *républicaine* or *de la République* ("[French] Republic schools", two and one occurrences, respectively), but also as *écoles de la mécréance* ("disbelievers' schools") and *écoles de la jâhiliyah* ("schools of ignorance or barbarism", two and six occurrences, respectively). Again, author-

itative Islamic texts are drawn upon to argue that the education system is not only flawed, but in fact corrupts the children's faith. The solution, according to the article, is *hijrah*, i.e. immigration to the caliphate, where the education programmes are "truly Islamic" (*réellement islamiques*) and "purified" of all disbelief and sins (*purifiés de toutes les mécréances et péchés*).

The category of "Reward" is included in Table 2 as, despite being less frequent than some of the other categories discussed in this paper, it is the one where the relative frequencies between English and French were closest together. The texts refer to several different types of reward, including Allah's blessings or mercy, but also the prospect of reward in the afterlife, sometimes referred to as the *hereafter* (84 occurrences). In the hereafter, the worthy will get to live in *Jannah*, the Arabic term for *paradise/paradis*. Paradise described as being achieved through dying while fighting the various "Others" described previously, and is thus strongly linked to martyrdom. However, attaining paradise is not the only reward referred to in the texts. More earthly and material rewards are also mentioned, referring to the literal "spoils of war" including *booty* (*butin*, 15 occurrences) (the Arabic *ghanīmah*) and *slaves* (*esclave* and *esclaves*, 20 and 22 occurrences, respectively), as shown in Examples (7) and (8).

(7)   Strike their shirk with your tawhīd, and Allah will break their strength. Allah will transfer their wealth and arms to you. You will seize it from their hands as war booty through which you will support Allah's religion and continue to be the people of support.

(8)   […]ensuite un cinquième de ces esclaves est transféré aux autorités de l'Etat Islamique pour être distribué en tant que cinquième du butin.
("[…]then a fifth of the slaves are transferred to the Islamic State to be distributed as a fifth of the booty")

Another word that we have classed as reward is *jizyah/jiziyah* which consists of tax levied by Islamic states on certain non-Muslim subjects. *Dar al-Islam* dedicates an entire section of Issue 5 to the superiority of ISIS in its application of the "judgment of Allah", for example, through the application of *jiziyah*. An illustrative instance is shown in Example (9).

(9)   L'Etat Islamique applique le jugement d'Allâh. L'Etat Islamique impose la Jiziyah des mécréants (Gens du Livre). contrairement aux Sahawât.
("The Islamic State applies the judgment of Allâh. The Islamic State imposes Jiziyah on disbelievers (People of the Book). Unlike the Sahawât.")

The imposition of jiziyah is just one example of how the Islamic State argues it follows a more rigorous and strict (i.e. "true") version of Islam; this, it is argued, is part of what distinguishes ISIS from other Islamic groups (e.g. *Sahawât*, i.e.

enemies or alternative militias argued by ISIS to be pro-USA), which are argued to be apostates. In other words, the imposition of *jiziyah* on disbelievers is part of ISIS's strategy of identifying as the most pure – and violent – of the Islamic groups (a finding also supported by previous research, e.g. Droogan & Peattie, 2016; Frampton et al., 2017; Ingram, 2017; Novenario, 2016). We now move on to consider categories that were much more frequently mentioned in the English texts as opposed to the French ones.

4.2   English categories

The English texts contained a much higher proportion of keywords related to the following categories: "Violence, war and death", "Places" and "Countries and travel". These categories and their associated words and frequencies are shown in Table 3.

Table 3. Categories more frequently referenced in the English texts

Category	Keywords in English corpus	English total frequency and relative frequency	Equivalents for French keywords
Violence, war and death	airstrikes, armies, army, assault, attack, attacks, battle, battles, blood, bomb, camp, camps, combat, crusade, crusaders, defending, destroy, enemy, enemies, factions, intelligence, defense, explosive, explosives, enmity, executed, fight, fighting, fighters, fitnah, forces, fought, front, fronts, imprisoned, jihad, jihād, jihadi, kill, killed, killing, military, mujahid, mujāhid, mujahideen, mujahidin, mujahidin, operation, operations, oppressed, oppression, raid, ribāt, security, soldiers, strike, sword, target, targeting, terror, terrorism, victory, wage, war, weapons, wires	15,184 20.23	2,377 12.42
Travelling, movement, journey	follow, forth, hijrah, muhājir, muhājirīn, path	1,052 1.54	168 0.87
Places	abyan, afghanistan, america, ayn, baghdādī, countries, dār, dārul, earth, egypt, turkey, dimashq, yarmūk, yemen, halab, idah, iran, khurāsān, land, lands, libya, makkah, mosul, tunisia, nation, nationalism, nationlist, pakistan, palestine, peninsula, raqqah, sham, sinai, somalia, state, syria, territory, west, western, wilāyah, wilāyat, wilāyāt	8,766 12.89	1,457 7.61

The English texts appear to be more focussed on violence and warfare compared to the French texts, with this category also containing the most keywords (both in terms of word types and overall tokens) of the keyword categories. In the English texts the most frequently occurring words were the Arabic word *jihad* (and its associated spellings occurring 1,367 times), followed by the verbs *fight* and *kill* (and related forms like *fought* and *killed*). *Jihad* generally translates as a struggle with a praiseworthy aim but has multiple interpretations. In a small number of cases, such as Example (10), this multiple sense of *jihad* is made explicit in the text.

(10)   Muslims have been legislated to wage Jihad for Allah's sake, to raise His word and make the Religion solely for Allah. This Jihad can take different forms, shapes and the use of different types of weapons to fight the enemy. And one of the weapons is striking and targeting the enemy's economic interests wherever they may be found.

However, in the English texts *jihad* is normally interpreted in the violent sense. For example, it collocates with the verb *wage* and related forms like *waging* 19 times. In the British National Corpus (a reference corpus which contains 100 million words of general British English which we accessed via the online tool *CQPweb*; Hardie 2012), the verb lemma WAGE collocates most strongly with the word *war*, with other strong (log-likelihood) collocates being *battle*, *warfare* and *campaign*. Thus the verb *wage* is primed in general British English (Hoey, 2005) to appear before the concept of warfare, indicating that this is likely to be the intended meaning of *jihad* within the extremist texts.

A second collocate of *jihad* is *path*, occurring 27 times and used in phrases like *path of jihad*. The concept of a path is likely to be familiar with Muslims, as it is used in a metaphorical way in texts that are part of the Islamic canon, e.g.:

> Those who would hinder (men) from the path of Allah and would seek in it something crooked: these were they who denied the Hereafter!
>
> (**Surah Hud 11:19, Quran**)

This structure is regularly found in the English corpus where *Allah* collocates with *path* 75 times and is one of its top ten collocates. Therefore, use of a similar metaphor, modified to refer to *jihad*, rather than Allah, creates equivalence between the two concepts within the corpus, legitimating the concept of violent jihad by recontextualising the wording of the more established text.

As well as *path*, the English texts have more words relating to journeys and travel more generally (a finding supported by the work of Droogan & Peattie, 2016), with three words from Arabic *hijrah* (196 occurrences), *muhājirīn* (54 occurrences) and *muhājir* (67 occurrences) also referenced. *Hijrah* (also *Hegira*) originally refers to the migration or journey of the Prophet Muhammad and His

followers from Mecca to Yathrib. It is associated with relocating to Muslim lands in order to escape persecution. In the corpus, the three strongest collocates of *hijrah* are *performed, performing* and *perform* (occurring 42 times collectively). In the BNC, the verb lemma PERFORM has *task(s)*, *duty/duties* and *service* (especially referring to *military service*) in its top 20 noun collocates (also using log-likelihood). The conceptualisation of *hijrah* as performance thus holds a priming for an obligatory duty, a point which is made clear by another of its collocates, *obligation* (see Example (11)).

(11)   Hijrah is an obligation from dārul-kufr [land of disbelief] to Dārul-Islām [land of Islam]

A related word *muhājirīn* (also *muhājir*) refers to the emigrants who went with Muhammad to Medina and now generally refers to Muslims who perform hijrah. As with *hijrah*, *muhājirīn* has been interpreted in the English texts to refer to people engaged in violent jihad.

The third category in Table 3 contains references to names of countries, cities and more general references to place. Such references are about twice as frequent (relative to the size of the corpora) in the English texts compared to the French ones, indicating that the English texts seem to be more focused on international events, particularly those occurring in countries which have significant Muslim populations like *Afghanistan*, *Egypt*, *Iran*, *Syria*, *Libya*, *Turkey*, *Pakistan*, *Palestine*, *Somalia* and *Yemen*. However, references to *America* and *western* also indicate interest in places where Muslims make up a significantly smaller proportion of the population. With the status of English as a global language, the audience for the English texts is perhaps expected to be wider than the audience for the French texts, which might explain why the English texts take more of an international perspective. Additionally, a key feature of the English texts (particularly *Dabiq* magazine) is the focus on exhorting readers to travel to Syria to engage in jihad while another is to inform (and inspire) readers about conflicts involving Muslims in a range of countries, as in Example (12).

(12)   Today American Muslims are being killed in Yemen, tomorrow they are gonna be killed in New York and Los Angeles.

A final aspect of the English texts, which can be seen from Tables 2 and 3, is the higher use of Arabic words. Taking the top 500 keywords, 140 of the English keywords and 77 French keywords were derived from Arabic. The strong prevalence of Arabic words in the English texts perhaps helps to give these texts a sense of authenticity (in the same way as some of the texts rely on reference to Islamic scripture) and requires readers who are not immediately familiar with Arabic to further engage with the extremist sources in order to find interpre-

tations, translations, and explanations (Frampton et al., 2017:25). This perhaps speaks to the different intended readerships: in English-speaking UK, most Muslims trace their lineage to Pakistan and Bangladesh (Ali, 2015:24), where Arabic is mainly encountered in religious scripture; in contrast, most Muslims in France come from Morocco, Algeria and Tunisia (Laurence & Vaïsse, 2006:15), where Arabic is widely spoken.

4.3 French categories

Although the English corpus seemed to be more focused on violence and warfare (see above), the French corpus, in fact, contained similar themes in different guises as shown in Table 4.

Table 4. Categories more frequently referenced in the French texts

Category	Keywords in French corpus	French total frequency and relative frequency	Equivalents for English keywords
Peace	armistice, hudnah, liberation, libérer, trêve, paix	396	589
		2.06	0.86
Obligation and laws	application, autorité, autorités, condition, décision, établir, interdit, jugement, législation, licite, mains, necessaire, ordonné, pacte, pouvez, règles, serment, traité, lois, obligation, obligatoire, permission, chari, charî	1,153	1,987
		6.02	2.92
Quote, say	cité, déclaré, demanda, parole, porté, rapporte, rapport, dirent, dis, dit	1,373	2453
		7.17	3.60
Islamic scripture, schools	livre, sounnah, sourate, hadith, hadith, boukhârî, bukhârî, coran, ṣaḥîḥ, sunnah, vers, verset, versets	939	1,169
		4.90	1.71
Emphasisers, discourse markers	afin, ainsi, auprès, certes, comme, consequent, delà, dépit, également, énorme, entièrement, general, malgré, même, mêmes, plutôt, quant, sûr	2,092	1,889
		10.9	2.77

For example, a closer investigation of the category "Peace", which had more word types with higher relative frequencies in French than in English, actually showed that peace was rejected by the groups. Most references to peace itself were used in fixed or semi-fixed honorific phrases pertaining to the Prophet Muhammad (e.g. *sur lui la prière et la paix; paix et bénédictions sur lui*). Although peace is occasionally called for as an ultimate goal of the creation of "worldwide jihad"

([…]*détruire le monde actuel et pour le remplacer par un nouveau/ ancien monde de justice et de paix*, "[…]destroy the present world and to replace it with a new/ old world of justice and peace"), truces (*hudnah, trêve,* 69 and 24 occurrences, respectively) were not seen as not desirable or permissible (Novenario, 2016). As shown in Example (13), it is argued that cease-fires should only allow fighters to catch their breath and then recommence their attacks on the "disbelievers".

(13)   Elle [la hudnah] n'est pas la manifestation de la recherche de la paix universelle à laquelle l'Islam aspirerait. Non ! Elle permet juste de reprendre son souffle pour mieux attaquer les mécréants
("It [hudnah] is not a show of universal peace to which Islam aspires. No! It only allows us to catch our breath in order to better attack the disbelievers")

Another category that is more predominant in French than in English is one pertaining to obligation and laws. Although previous research on English-language extremist texts found evidence of dominant themes relating to obligation and jurisprudential justifications (e.g. Droogan & Peattie, 2016; Frampton et al., 2017; Ingram, 2017), the theme was found to be far more salient in the French texts. The array of word types referring to obligation is diverse, ranging from specific references to Sharia law (*chari, charî,* 33 and 70 occurrences, respectively) to general references to laws, legislation, treaties, pacts and rules (*lois,* 100; *législation,* 24; *traités,* 8; *pacte,* 172; *règles,* 51). These tend to be described as deriving from Allah (e.g. *la législation divine/islamique/d'Allâh).* There are also discussions of things being permitted (*licite,* 26) and forbidden (*interdit,* 58), as well as references to often unspecific sources of authority, decisions, obligation and permission, which are more open to interpretation (*autorité,* 50; *autorités,* 17; *décision,* 18; *obligation,* 50; *obligatoire,* 46; *permission,* 42). There are also metaphorical invocations of power with relation to hands (e.g. *Allah, par vos mains,* "Allah, by your hands"; *par leurs proper mains,* "with their own hands").

As noted above, the references to obligation seem to emerge because ISIS self-represents as imposing Islamic law "better" (i.e. more thoroughly) than its competitors, as shown in Example (14).

(14)   […]d'autres factions qui refusent l'application des préceptes de l'Islam et qui font partie de l'alliance des Ṣaḥawât.
("[…]other factions that refuse the application of Islamic precepts and that are part of the Sahawat alliance.")

ISIS advocates "applied theology" (Frampton et al., 2017) that supports calls for action rather than "just" belief, and it is through their thorough (violent) imposition of a strict (mis)reading of Islamic law that ISIS claims its own authority over

all other Muslim groups, including countries where Sharia law is not in effect (see Example (15)).

(15)   Il ne s'agit pas là de digresser sur la mécréance des gouverneurs qui n'appliquent pas la Charia car cette question fait partie des axiomes de la religion musulmane.
("We won't digress here onto the subject of infidel governors who do not apply Sharia because this is a pillar of Islamic faith.")

The range of topics referred to with relation to obligation is not overly diverse. The overarching obligation is to establish (*établir*) a land of Islam, where Sharia law (the law of Allah) dominates. Generally readers are called upon to follow "true" Islam, reject "corrupt" society, emigrate to the caliphate and/or take up arms. Although these obligations were also themes in the English corpus and pertain to findings from previous research (Novenario, 2016), they were more salient in the French texts, where they are discussed with relation to a higher power. For example, taking up arms is repeatedly discussed with relation to being "ordered to fight" (*ordonné de combattre*, 7 occurrences). The legitimacy of polygamy is another theme; in Issue 7 of *Dar al Islam*, women are reminded that their husbands have no *obligation* to consult them or to ask for their permission (*authorisation*) or approval (*agrément*). As shown in Example (16), women are advised to submit to Allah and obey His law, which permits polygamy.

(16)   En outre, chaque sœur doit savoir que si son mari désire se remarier, rien ne lui impose de la consulter, ni de lui demander l'autorisation, ni de rechercher son agrément.
("Also, each sister should know that if her husband wishes to remarry, nothing requires him to consult her, ask her permission, or to seek her approval.")

Given the importance of obligation deriving from authorities and authoritative texts in the French magazines, it follows that other categories unique to the French dataset pertain to quotation and Islamic scripture. While this is also the case in the English corpus (and in previous research, e.g. Droogan & Peattie, 2016), the theme is far more salient in French. Quotation (e.g. *cite*, 28 occurrences) and statement (*déclaré*, 25; *rapporte*, 88; *dirent*, 34; *dis*, 74; *dit*, 767) verbs refer almost exclusively to the source texts used to substantiate potentially contentious arguments made by the authors (i.e. arguments that are counter to established western ways of thinking, such as the advocation of violence). Source texts are often referred to with relation to the author (e.g. *boukhârî/bukhârî*, 50 and 41 occurrences, respectively) or to the name, unit, or type of the text itself (e.g. *sounnah*, 59; *sunnah,* 24; *sourate*, 143; *hadith*, 106; *coran*, 40; *vers*, 166; *verset*, 184; *versets*, 50).

The largest category in French compared to English pertains to emphasisers and discourse markers. These are not unusual words in written French, but they emerged as statistically significant in comparison with the corpus of *La Presse* newspapers. It is possible that the unusually frequent use of emphasisers and discourse markers is the result of the magazines employing a highly formal register; this might have been an intentional stylistic choice to give the impression of authority and intelligence. However, this might also have been the result of the comparison of two different genres (the magazine format of the French extremist texts vs. the French newspapers used as the reference corpus) or perhaps the comparison of two different varieties: *La Presse* is a French Canadian publication, whereas *Dar al-Islam* is focused on France and its authors seem to employ European French. The availability of a larger, general reference corpus of French would have helped us to ascertain whether this category of words was, in fact, unique to extremist discourse.

5.   Conclusion

Our analysis shows that the English and French texts contain many similarities. One core focus in both datasets was religion, a theme in extremist discourse widely misunderstood in the West (Frampton et al., 2017). Calls for violence are consistently made based on purported devotion to and obligation deriving from divine authority; these are unfailingly supported by (mis)readings of authoritative religious figures or texts. A strong focus on othering occurs in both sets too (a finding supported by previous research, e.g. Droogan & Peattie, 2016; Ingram, 2017; Novenario, 2016), along with reference to reward – reward in paradise for being a martyr (dying while committing violence) but also the promise of more earthly rewards such as *booty*.

There are, however, important differences between the English and French extremist discourse. The English texts appear to be more concerned with world events pertaining to Islam, which may reflect the expectation of a more global audience. The French texts focus particularly on aspects of French culture, such as schools in France, and also seem more likely to use quotation from scripture to not only legitimate their perspectives, but also to "educate" the reader (Frampton et al., 2017). In contrast, the English texts are more likely to code-switch to Arabic terms as a form of legitimation. A possible reason for this may be due to the discourse communities associated with English-speaking and French-speaking extremist Islamists. As noted above, in France, many Muslims come from Morocco, Algeria and Tunisia (Laurence & Vaïsse, 2006:15) where Arabic is widely spoken. In English-speaking UK, Muslims tend to have originated from

Pakistan and Bangladesh (Ali, 2015: 24) and are less likely to be familiar with Arabic, apart from encountering it in religious scripture, so the use of Arabic in the English texts may function as a way of priming readers to view the texts as authentically religious. There is also the possibility that readers, accustomed to the international status of English, are being initiated into a world where Arabic is the common language, shared both by seasoned militants and new recruits. The propaganda is therefore being used to bridge language gaps (Frampton et al., 2017).

The French texts have a stronger focus on permissions, rights, obligations and laws and employ a more formal register concomitant with this focus. In contrast, the English texts seem more concerned with advocating violence and inspiring readers to engage in jihad at home or to travel to distant lands for the same purpose. This difference might have arisen due to the mixed sources of the English texts. *Dabiq* has been found to be more formal than *Inspire* in its language, with the latter magazine more renowned for its colloquial English, Americanisms, and popular appeal (Ingram, 2017), and the intermixing of these texts in the corpus might have diluted the salience of formal language. When the French texts refer to violence, it is more likely to be viewed as a negation of peace or justified as a means to eventual peace – a legitimation strategy not seen to the same extent in the English texts.

Methodologically, the examination of cross-linguistic categories rather than words allowed us to examine broader trends and the ways in which meanings can be expressed in a range of different ways in both languages. This has allowed us to consider additional linguistic and cultural perspectives on the issue of terrorism, which has important global relevance and currency. Such an approach might prove to be useful in the study of a range of other global topics, such as climate change and the international rise of populism. In our study, due to the nature of the text types, we found it advantageous to manually create categories, rather than relying on pre-existing and automatic categorisation schemes; this might also be the case in studies of other issues with other text types. We acknowledge the difficulty in comparing two sets of texts written in different languages and note that the English corpus was drawn from a wider range of sources and was also larger than the French corpus; the small size of the French corpus in particular could have led to individual articles skewing the findings (e.g. most of the references to *école/s* in the French corpus came from a single lengthy article). We also recognise that our own reliance on English-medium literature and research in the writing of this paper could have potentially biased our exposure to alternatives available in other languages. These issues highlight a need for more work on corpus creation and analysis both in French itself and in experimental studies that compare different languages. More generally, the analysis suggests that counter-

terrorism messages need to be tailored for different language speakers who would have encountered different strategies in the original messages.

References

Aggarwal, N. K. (2017). Exploiting the Islamic State-Taliban rivalry for counterterrorism messaging. *Journal of Policing, Intelligence and Counter Terrorism*, 12(1), 1–15. https://doi.org/10.1080/18335330.2016.1223868

Ali, S. (2015). *British Muslims in Numbers*. London: The Muslim Council of Britain.

Anthony, L. (2016). *AntConc* (Version 3.4.4) [Computer software]. Tokyo: Waseda University.

Baker, P. (2006). *Using Corpora in Discourse Analysis*. London: Continuum.

Baker, P., Gabrielatos, C., & McEnery, T. (2013). *Discourse Analysis and Media Attitudes: The Representation of Islam in the British Press*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511920103

Buril, F. (2017). Changing God's expectations and women's consequent behaviors – How ISIS manipulates "Divine Commandments" to influence women's role in Jihad. *Journal of Terrorism Research*, 8(3), 1–10. https://doi.org/10.15664/jtr.1363

Conoscenti, M. (2016). ISIS' *Dabiq* communicative strategies, NATO and Europe. Who is learning from whom? In M. Ceretta & B. Curli (Eds.), *Discourses and Counter-discourses on Europe: From the Enlightenment to the EU* (pp. 215–244). London: Routledge.

Droogan, J., & Peattie, S. (2016). Reading jihad: Mapping the shifting themes of *Inspire* magazine. *Terrorism and Political Violence*, 30(4), 684–717. https://doi.org/10.1080/09546553.2016.1211527

Esposito, J. L. (2014). *The Oxford Dictionary of Islam*. Oxford: Oxford University Press.

Frampton, M., Fisher, A., & Prucha, N. (2017). *The new netwar: Countering extremism online*. London: Policy Exchange. Retrieved from: https://policyexchange.org.uk/publication/the-new-netwar-countering-extremism-online/ (last accessed August 2018).

Gabrielatos, C., & Marchi, A. (2012, September) *Keyness: Appropriate metrics and practical issues*. Paper presented at CADS 2012, Bologna, Italy.

Ghilan, M. (2013, June 8). It's extremist Muslims, not Islamic extremism. *Al Jazeera*. Retrieved from http://www.aljazeera.com/indepth/opinion/2013/06/2013681251485552.html (last accessed August 2018).

HM Government. (2013). *Tackling extremism in the UK*. Cabinet Office: London. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/263181/ETF_FINAL.pdf (last accessed August 2018).

Hoey, M. (2005). *Lexical Priming: A New Theory of Words and Language*. London: Routledge. https://doi.org/10.4324/9780203327630

Ingram, H. J. (2017). An analysis of *Inspire* and *Dabiq*: Lessons from AQAP and Islamic State's propaganda war. *Studies in Conflict & Terrorism*, 40(5), 357–375. https://doi.org/10.1080/1057610X.2016.1212551

Kramer, M. (2003). Coming to terms, Fundamentalists or Islamists? *Middle East Quarterly*, (Spring 2003), 65–77.

Laurence, J., & Vaïsse, J. (2006). *Integrating Islam: Political and Religious Challenges in Contemporary France*. Washington, D.C.: Brookings Institution Press.

Novenario, C. M. I. (2016). Differentiating Al Qaeda and the Islamic State Through Strategies Publicized in Jihadist Magazines. *Studies in Conflict & Terrorism*, 39(11), 953–967. https://doi.org/10.1080/1057610X.2016.1151679

O'Halloran, K., Tan, S., Wignell, P., Bateman, J. A., Duc-Son, P., Grossman, M., & Vande Moere, A. (2016). Interpreting text and image relations in violent extremist discourse: A mixed methods approach for big data analytics. *Terrorism and Political Violence*. Advance online publication. https://doi.org/10.1080/09546553.2016.1233871

Partington, A. (2004). Corpora and discourse, a most congruous beast. In A. Partington, J. Morley & L. Haarman (Eds.), *Corpora and Discourse* (pp. 11–20). Bern: Peter Lang.

Prentice, S., Taylor, P., Rayson, P., Hoskins, A., & O'Loughlin, B. (2011). Analyzing the semantic content and persuasive composition of extremist media: A case study of texts produced during the Gaza conflict. *Information Systems Frontiers*, 13(1), 61–73. https://doi.org/10.1007/s10796-010-9272-y

Prentice, S., Taylor, P., Rayson, R., & Giebels, E. (2012). Differentiating act from ideology: Evidence from messages for and against violent extremism. *Negotiation and Conflict Management Research*, 5(3), 289–306. https://doi.org/10.1111/j.1750-4716.2012.00103.x

Rayson, P. (2009). *Wmatrix: A Web-based Corpus Processing Environment* [Computer software]. Lancaster: Lancaster University.

Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam/Philadelphia, PA: John Benjamins. https://doi.org/10.1075/scl.6

Vergani, M., & Bliuc, A. M. (2015). The evolution of ISIS' language: A quantitative analysis of the language of the first year of Dabiq magazine. *Sicurezza, Terrorismo e Società*, 2, 7–20.

Vessey, R. (2013). Challenges in cross-linguistic corpus-assisted discourse studies. *Corpora*, 8(1), 1–26. https://doi.org/10.3366/cor.2013.0032

Wignell, P., Tan, S., & O'Halloran, K. (2017a). Under the shade of AK47s: A multimodal approach to violent extremist recruitment strategies for foreign fighters. *Critical Studies on Terrorism*, 10(3), 429–452. https://doi.org/10.1080/17539153.2017.1319319

Wignell, P., Tan, S., & O'Halloran, K. (2017b). Violent extremism and iconisation: Commanding good and forbidding evil? *Critical Discourse Studies*, 14(1), 1–22. https://doi.org/10.1080/17405904.2016.1250652

Wignell, P., Tan, S., O'Halloran, K., & Lange, R. (2017c). A mixed methods empirical examination of changes in emphasis and style in the extremist magazines Dabiq and Rumiyah. *Perspectives on Terrorism*, 11(2), 2–20.

Zelin, A. Y. (2015). Picture or it didn't happen: A snapshot of the Islamic State's official media output. *Perspectives on Terrorism*, 9(4).

Address for correspondence

Paul Baker
Department of Linguistics and English Language
Lancaster University
County South
Lancaster, LA1 4YLUK
UK

j.p.baker@lancaster.ac.uk

Co-author information

Rachelle Vessey
Department of Applied Linguistics and
Communication
Birkbeck College, University of London
r.vessey@bbk.ac.uk