Representing Religion in the British Press: A Corpus-based Analysis

Sawsan Hassan (University of Huddersfield, UK)

Introduction

The study is concerned with investigating religion in all British broadsheet newspapers in two periods of time (January 2001 and January 2010). The purpose of choosing such periods is because of the global awareness of religion that followed the events occurred in the first decade of 21st century. Recent studies (Weller, 2011) have shown the place and the impact of religion on the public in Britain in the first decade of this century because of a number of events such as the New York 9/11 attacks of 2001, the Madrid bombing 11/3 of 2004, the London 7/7 bombing of 2005, and the Glasgow 30/6 bombing of 2007, as well as the invasions of Afghanistan in 2001and Iraq in 2003 and other events. Thus religion has radically changed and has broken into mainstream Western news agendas (Gower and Mitchell, 2012).

I intend to analyse the press representation of three widespread religions in the UK – Christianity, Judaism, and Islam as it is impossible to cover all world religions in one research study. These three monotheistic religions are mostly practiced in the UK according to the census data in 2011 in the UK (ONS 2011); Christianity occupied the majority; Islam is next, and then Judaism.

Investigating religions has been the focus of a range of frameworks in different disciplines such as media, theology, and sociology, but it has not been tackled with within the area of linguistics. Accordingly, this study is conducted using a selected set of corpus linguistic tools namely frequency lists, concordances, and collocations.

I present a case study investigating applicability of the corpus tool Sketch Engine to sort collocates of those three religions. I aim to examine language used to represent these religions and their followers as well, i.e. how many times a certain group have been represented in the press as grammatical agents or patients in order to illustrate any changes in representation of religion over time.

Data collection

The data is comprised of news articles from British broadsheets. It covers two periods: January 2001 and January 2010. All the broadsheet newspapers and their Sunday versions were selected: The Guardian, The Observer, The Independent, The Independent on Sunday, The Daily Telegraph, The Sunday Telegraph, The Times, The Sunday Times, The Financial Times, and The Business. Using ProQuest database, I ran a search for the following query terms: Christianity! OR Christian! Islam! OR Muslim! OR Islamic! Judaism! OR Jew! OR Jewish!

The data consists of 1032 and 1168 articles in 2001 and 2010 respectively after excluding all duplicated and unrelated articles. All articles were then uploaded into Sketch Engine to conduct a corpus analysis. The number of words in each corpus is 797,336 in 2001 and 994,081 in 2010.

As far as these three Abrahamic religions are concerned, I intend to view the raw frequency (RF), which needs to be normalised based on the normalised frequency (NF) per 1.000 words as the size of the two corpora is different. McEnery et al. (2006: 52-3) state that when comparing two corpora of different sizes, raw frequencies extracted from corpora need to be normalised to a common base. The following table show the raw and normalised frequencies for each search query:

Words	2001/RF	NF	2010/RF	NF
Christianity	79	0.09	92	0.09
Christians	87	0.10	162	0.16
Christian	537	0.67	644	0.64
Judaism	48	0.06	25	0.02
Jews	406	0.50	240	0.20
Jew	253	0.25	149	0.14
Jewish	677	0.84	446	0.44
Islam	64	0.08	229	0.23
Muslim	295	0.36	883	0.88
Muslims	123	0.15	380	0.38
Islamic	231	0.28	453	0.45

Table (1): The raw and normalised frequencies of the search items

From the table, it is apparent that the frequency of the search items is varied in both periods of time in which some decreased, while others increased in 2010. For instance, 'Judaism' is the least search word in both periods of time; 'Christianity' kept the same level of frequency, while 'Islam' increased considerably (0.08 to 0.23). It is worth mentioning that the singular forms of some search words in both corpora tend to occur as a noun and an adjective as in the case of Christian and Muslim. As a result, the singular form of Christian and Muslim in both periods is higher than its plural form as shown in the table above, but the adjectival pattern Jewish, is declined sharply in 2010 (0.84 to 0.44). words 2001/RF NF 2010/RF NF Christianity 79 0.09 92 0.09 Christians 87 0.10 162 0.16 Christian 537 0.67 644 0.64 Judaism 48 0.06 25 0.02 Jews 406 0.50 240 0.20 Jew 253 0.25 149 0.14 Jewish 677 0.84 446 0.44 Islam 64 0.08 229 0.23 Muslim 295 0.36 883 0.88 Muslims 123 0.15 380 0.38 Islamic 231 0.28 453 0.45

Regarding the followers or believers of these three religions, in the first place the singular form of Christian increased in 2010 in terms of raw frequency 537 to 644, but decreased slightly in terms of normalised frequency (0.67) to (0.64) though the size of the corpus in 2010 is larger than that in 2001. However, the plural from Christians rose from 87 (0.10) to 162 (0.16). In addition, Muslim trebled in 2010 (0.36 to 0.88), and its plural form Muslims doubled (0.15 to 0.38). Jew or Jews decreased in 2010 in terms of both raw and normalised frequencies.

Word sketches

Having outlined the frequency differences of the items under investigation, I intend to report the findings of word sketches that I carried out for each search word. I uploaded my data into Sketch Engine – a corpus query tool developed by Kilgarriff et

al. (2004) to identify collocates in a range of grammatical relations via Word Sketch. Sketch Engine is an important and useful tool to analyse the grammatical features of salient collocates. It not only identifies collocates, but also specifies the grammatical constructions within such collocates (Baker et al. 2013: 37). Therefore I used Sketch Engine to create a word sketch for each search query. Word Sketch is a one-page automatic, corpus-based summary of a word's grammatical and collocational behaviour, (Kilgariff et al. 2004). It provides a summary of the main part of speech of collocates of the target word.

Creating a word sketch for each search word, i.e. Christianity, Christian, Judaism, Jew, Jews, Jewish, Islam, Muslim, Muslims, and Islamic, I investigated the grammatical relations between collocates which can tell me something about phenomena like agency. Word Sketch provides a set of significant collocates of each item sorted into various grammatical frames or patterns. For instance, verbs and adjectives co-occur with the search query might reveal how a certain group such as Christians, Jews, and Muslims, are presented as agents or goals, and how they are described in the British newspapers. This therefore leads me to shed some more light on the nature of agency extracted from the sorts of verbs collocating with the search items. Further Christian, Jewish, Muslim, and Islamic, as adjectives, can identify the nouns that are collocating with them, and so reveal which sorts of nouns such items describe. Therefore, a special attention is paid to three grammatical patterns: subjects, objects, and modifiers.

The following table shows the number of occurrences of each item in the word sketch plus its statistics in both periods. The statistic score is measured by logDice based on how salient the words are. Salience is the basic measure of Word Sketch. Moreover, LogDice is based on Dice coefficient which reveals the more frequent lexical collocates (For more information, see Rychlý, 2008).

Items	On. of occurrences/2001	Statistics pm	On. Of occurrences/2010	Statistics pm
Christianity	79	82.48	92	76.95
Christian (n.)	282	294.43	311	260.15
Christian (adj.)	255	266.24	333	278.55
Judaism	48	50.11	25	20.19
Jew	250	261.02	149	124.64
Jews	233	243.27	140	117.11
Jewish	635	663.00	421	352.17
Islam	64	66.82	229	191.56
Muslim (n.)	119	124.24	344	287.76
Muslim (adj.)	179	183.76	539	450.88
Muslims	72	75.17	197	164.79
Islamic	179	186.89	325	272.16

Table (2): Number of occurrences of the search items
in the word sketch in 2001 and 2010

In terms of statistics, there is a decline in the findings regarding Judaism, Jew (s), and Jewish, while it is the opposite in the case of Islam, Muslim (s), and Islamic in 2010. Christianity and Christian kept almost the same statistic scores in both periods.

By way of example, Jews decreased from (243.27) in 2001 to (117.11) in 2010, Muslims increased from (75.17) in 2001 to (164.79) in 2010. Further a word sketch viewed that the result of Christian as a noun and an adjective is somewhat distinct (294.43) and (266.24) respectively. But Jewish is more frequent and salient than the noun Jew (s) (table 2). Similarly, Muslim tended to be used as an adjective (183.76) more frequently than as a noun (124.24).

It is worth mentioning that word sketches of Christian, Jew, and Muslim tagged as singular nouns display also their plural forms; thus I consider the plural forms alone to find out how such groups (Jews and Muslims) are represented and behave. However, the sketch has not shown any significant collocates in the grammatical features in the case of Christians in both periods; only 3 (3.13 per million) and 6 (5.01 per million) in 2001 and 2010 respectively. As a result, I examined Christians in connection with the singular form as the word sketch encompasses both forms of the target word.

References

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

Gower, O. & Mitchell, J. (2013). *Religion and the News*. London: Routledge.

Kilgarriff, A., Rychly, P., Smrz, P. & Tugwell, D. (2004). Itri-04-08 the sketch engine. *Information Technology*, 105, 116.

Rychly, P. (2008). A Lexicographer-Friendly Association Score. In RASLAN 2008. 2. vyd. Brno, RASLAN 2008. Brno: Masarykova Univerzita, 2008. s. 6-9, 4 s. ISBN 978-80-210-4741-9.

Stubbs, M. (2001). *Word and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.

Weller, P. (2011). Religious discrimination in Britain: A Review of research evidence, 2000-10. *EHRC Research report*, 73. Manchester: Equality and Human Rights Commission.