

Edge Hill Corpus Research Group, 2 March 2023

A lone wolf from the ISIS pack: Hunting discourses through manual annotation

Daniel Malone

PhD Researcher

Edge Hill University

Abstract

The process of manual corpus annotation, where researchers add interpretive information to corpus data, is a valuable tool for systematically analysing linguistic or semantic features in a corpus. A core aspect of manual annotation is the *annotation scheme* – a set of guidelines for labelling corpus content, including annotation categories, definitions, and examples.

In this talk, I will introduce the manual annotation approach and annotation scheme I am developing to aid a critical discourse analysis of the representations of the nexus between lone-wolf terrorists and the extremist groups ISIS and al Qaeda in the UK press. The underpinning goal of this annotation scheme is to systematically reveal discourse prosodies, in other words, the implicit and explicit attitudes (Stubbs 2001: 66) towards the lone-wolf terrorist.

The context for this talk is my doctoral research project "Constructing the Lone Wolf Terrorist: A corpus-based critical discourse analysis." The dataset used in this study is The Lone Wolf Corpus, a purpose-built corpus consisting of approximately 8.5 million words and 8,600 texts from UK national newspapers, published between 2000 and 2019 (Malone 2020).

I will describe my analytical approach and the process I used to develop the annotation scheme, which involved iterative cycles of annotation, and explain the reasons for employing this approach. I will provide an overview of the scheme's five tags and illustrate the five values of *Tag 4: Degree of Connection* with detailed examples. Finally, I will address two practical challenges that arose during the development of the annotation scheme, namely the need to avoid top-down categorisation, and the difficulty of balancing scheme richness with labour efficiency.

Talk Structure

- The Project & Corpus
- The Approach
- Annotation Scheme*
- Detailed examples of *Tag 4: Connection**
- Reflection on problems encountered

*Works in progress

The Project

Objective: a critical discourse study of representations of the lone-wolf terrorist and surrounding discourses

Means: corpus approaches and techniques

Research Questions:

- RQ1-How is the lone-wolf terrorist represented by the British press?
- RQ2-(How) do representations of the lone-wolf terrorist change?
- RQ3-How do these portrayals contribute to surrounding discourses?

Texts: Newspaper articles from 31 UK Newspapers

Period: 20 years (2000-2019)

Database: Lexis Nexis

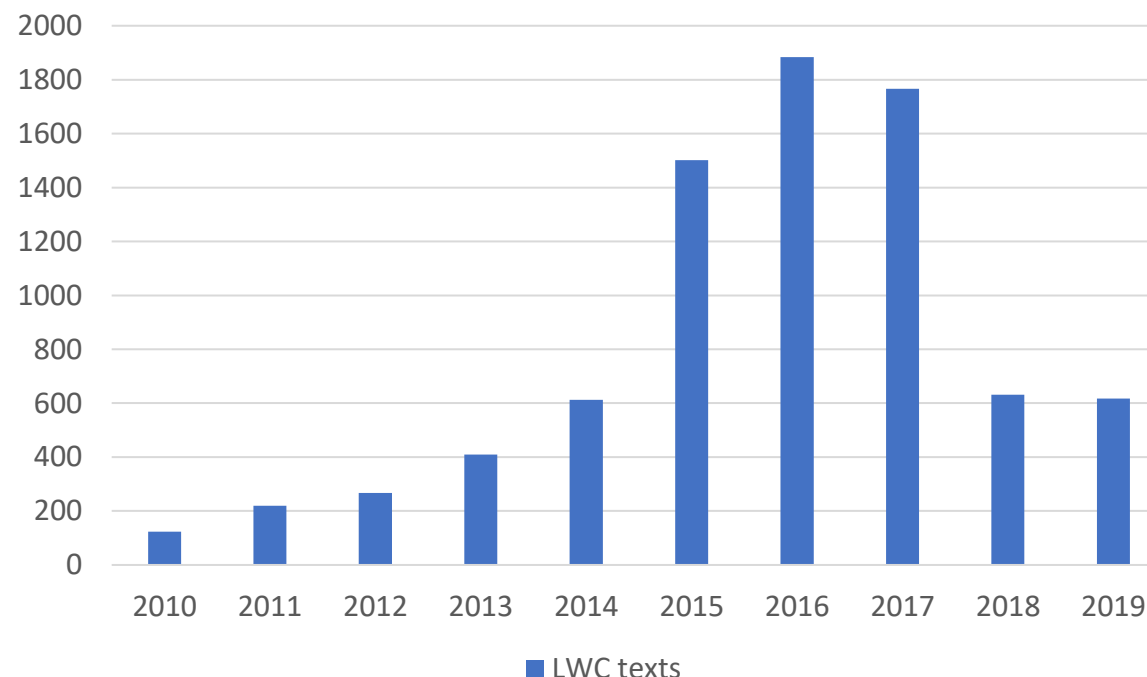
Query: Dual Group Query of 70 query-terms (Malone 2020) developed using an adapted version of the RQTR method (Gabrielatos 2007)

Subcorpora:

- Annual
- Pre-peak (2010-2014)
- Peak (2015-2016)
- Post-peak (2018-2019)

Lone Wolf Corpus

LWC annual text frequency 2010-2019



Contextual note: ISIS announced the establishment of a caliphate – the Islamic State – in mid-2014

The Approach

1. Collocation analysis - co-occurrences of two words that are unusually frequent (Sinclair 1991)
 - Node: *lone wolf, lone-wolf, lone wolves*
2. Semantic preference analysis - collocates examined to identify the potential associations of the node with groups of semantically-related words (Stubbs 2001: 65)
 - ISIS group: *Daesh, IS, I.S., ISIL, ISIS, Islamic State*
 - Al Qaeda group: *al Qaeda, al-Qaeda, al Qaida, al-Qaida, al-Qa'ida*
3. Discourse prosody analysis – reveal explicit and implicit attitudes (Stubbs 2001: 66) towards the lone-wolf terrorist
 - Extracted random samples of collocations from each subcorpora
 - Conducted close-reading of *extended* concordances - i.e., 75+ words either side of node
 - Employed a bottom-up approach to annotation - i.e., from a naïve position (as far as is possible)
 - Conducted iterations of annotation => creation and refinement of annotation scheme
 - Informed by van Leeuwen's Social Actor approach (2008)

Why this approach?



Improved identification of discourse prosodies and understanding of pragmatic nuance in the extended co-text



More likely to reveal inter-related discourses due to the extended co-text considered



Avoids “eye-balling” (Gillings & Hardie 2022) - i.e., judging discourses and pragmatic meaning based on *far-reading* - e.g., using collocate lists or restricted concordance lines with little co-text

Annotation Scheme

Tag 1: Node

- *What type of entity is described by the node or the noun compound involving the node?*

Tag 2: Collocate

- *What type of entity is described by the node or the noun compound involving the node?*

Tag 3: Event

- *What type of events link the node (...) and the collocate(...)?*

Tag 4: Connection

- *To what degree are the LW and EO connected in the representation?*

Tag 5: Discourses

- *What are the surrounding discourses?*

Tag 4: Degree of Connection

To what degree are the lone wolf and the extremist organisation connected in the representation?

Value 1: Distinct	Lone wolf (LW) and extremist organization (EO) are represented as distinct with no direct connection between them. LW is fully autonomous.
Value 2: Inspired	EO-related motivation is represented but without indication of direction, instruction or training.
Value 3: Informed	EO directs LW, directly or indirectly, with instructions, knowledge or suggestions/advice.
Value 4: Membership	Current or prior EO membership of LW is represented.
Value 5: Blurred	LW and EO are indistinct or merged in the representation.

Value 1: Distinct

... The **main threat** to the Olympics are **lone wolf terrorists**, MI5 said yesterday. **Al Qaeda sympathisers** and dissident Irish republicans may also pose a **potential threat**. Defence expert Will Hartley, head of the IHS Jane's Terrorism and Insurgency Centre, said: "The **threat** posed by **so-called 'lone wolf' terrorists- who act independently rather than as part of a militant group-** is of particular **concern** to security services because the **threat** is **so hard to quantify**." Police and intelligence officials are understood to have picked up an increase in chatter, **but there is not thought to be any intelligence of a credible plot to attack the Games. All leave has been suspended at MI5. ...**

[Mirror_2012_LW (3).txt]

Key & Discourses (e.g., C5)

Node (rest of noun)

Collocate (rest of noun)

LW synonym/definition

Topos of Fear/Threat

Discourse of Assurance

V1: Distinct

... RUSI identified **four key types of threat** to the Games: **an organised al-Qaeda plot**, a **lone-wolf attack**, disruption by Irish Republican dissidents or an attempt to assassinate any of the 120 or so heads of state who were due to attend. While protecting the Olympics was a significant and vital challenge, however, the think-tank stressed that **threats** to national security would still exist beyond the closing ceremony, **requiring sufficient funding even during a time of austerity**. ...

[TimesOnline_2012_LW (3).txt]

Key & Notes

Node (rest of noun)

Collocate (rest of noun)

Topos of Fear/Threat

Discourse of Action

V2: Inspired

... As a result of the shooting, **security was beefed up across the country at military recruitment centers and bases. The NYPD has also increased their presence at 'sensitive areas'- including recruitment centers and landmarks.** The attack came on **the last day of the holy Muslim festival of Ramadan, a period during which ISIS agitators** asked for more **lone wolf attacks. He was a practicing Muslim,** and would take breaks from training as a Mixed Martial Arts fighter to pray in his coach's office. **Fellow worshipers had seen him in a local mosque up to a month before the attack, but it is not known if or how he may have been radicalized.** ...

[DailyMailOnline_2015_LW (242).txt]

Key & Notes

Node (rest of noun)

Collocate (rest of noun)

Discourse of Assurance

Other (potential) motivation

V3: Informed

... Police have launched a terror investigation into the attack, which comes after calls by Isis and **al-Qaeda** for **supporters to launch "lone wolf" attacks in countries perceived as enemies**. The arrested man was locally identified as a 38-year-old construction worker from Karaman province, who was said to be **mentally ill**. Both **Kurdish separatists** and Islamist extremists have targeted Turkey with recent attacks, including a rocket and shooting attack on a police station in the province of Van on Monday. ...

[Independent_2016_LW (81).txt]

Key & Notes

Node (rest of noun)

Collocate (rest of noun)

LW synonym

Other (potential) motivation

V3: Informed

... Claimed ISIS supporters have been discussing using Ebola in attacks.

Militants also said to be **plotting** to use other **deadly** toxins and viruses.

ISIS spokesman has also used the internet to encourage **would-be jihadis**.

Suggested **supporters** carry out **lone wolf attacks** using poisoned needles.

ISIS is urging **supporters in the West** to **kill civilians by any means**

possible- including using Ebola as a chemical weapon and stabbing people

with poisoned needles, a senior Spanish politician claims. **Francisco**

Martinez, Spain's state secretary for security, said he is taking the threats

seriously...

[DailyMailOnline_2014_LW (74).txt]

Key & Notes

Node (rest of noun)

Collocate (rest of noun)

LW synonym

Topos of Threat

Discourse of Assurance

V4: Member

MANY IRISH JIHADIS DEAD IN WAR ON **ISIS**; Anti-terror chiefs **fear**
lone wolf attack by **fighters returning to Ireland**

MOST of **the 30 Irish jihadi fighters who joined ISIS** are believed to have been killed, a top anti-terror officer said yesterday. Assistant Commissioner Michael O'Sullivan, in charge of security and intelligence, was speaking at the Interpol European Regional Conference in Dublin. He told the Irish Mirror his officers most **fear** a terror massacre from **a lone wolf attacker**. **Mr O'Sullivan added he has tested the Garda response and believes the force will be able to deal with any threat.**

[Mirror_2018_LW (9).txt]

Key & Notes

Node (rest of noun)

Collocate (rest of noun)

Topos of Fear/Threat

LW Synonyms

Discourse of Assurance

V4: Member

... Air strikes are being unleashed by the US and the RAF. Military chiefs in Britain **fear the IS threat** will not end when it loses Raqqa and Mosul. A security source said: "With the fall of their geographical position **IS is likely to revert to more classic terrorist models.** "It is likely they will try to disappear underground like **al-Qaeda** and encourage **lone wolf attacks.**" Referring to **fanatics fighting now**, the source added: "They are **extremely dangerous,** skilled and hardened by years of battle with the know-how to cause **carnage on Europe's streets.**"

[Mirror_2017_LW (21).txt]

Key & Notes

Node (rest of noun)

Collocate (rest of noun)

Topos of Fear/Threat

Discourse of Preparedness

V5: Blurred

US **fears** 'lone wolf' attacks

AMERICA is **bracing itself** for a terrorist attack by an **Al Qaeda "lone wolf"** within the next three months. Intelligence chiefs **fear** the extremists already have **hidden agents**, or "**sleepers**", living in the US ready to carry out an atrocity. The **Al Qaeda suspects** sent to the US include what are known as "**clean skins**" - **recruits with few or no known links to extremists**. Central Intelligence Agency director Leon Panetta said the "**lone wolf**" **operators** were being inspired to plan and carry out their own attacks. ...

[Express_2010_LW (5).txt]

Key & Notes

Node (rest of noun)

Collocate (rest of noun)

LW Synonyms

Topos of Fear/Threat

V5: Blurred

..."Daesh have anti-tank weapons," says General Yassin, "and night vision binoculars. We don't even have those, which makes night attacks difficult. God willing we will prevail. **But what we need is help from the international community; equipment, supplies. We are protecting Europe. So where is Europe for us?"**

The report by Europol found that in 2015, **six** EU nations- Denmark, France, Greece, Italy, Spain and the UK- bore the brunt of **completed or failed IS attacks**, either carried out by **foreign lone wolf jihadis** or inspired by **IS**. Of the **1,077** arrests made, France had held **424** suspects, Spain **187** and the UK **134**. In the most recent large-scale attack, **84** people were killed and hundreds injured when French-Tunisian Mohamed Lahouaiej Bouhlel ploughed his lorry into crowds in the French Riviera city of Nice on July 14.

[ExpressOnline_2016_LW (116).txt]

Key & Notes

Node (rest of noun)

Collocate (rest of noun)

LW Synonyms

Topos of Number

Discourse of Action

What next?

More annotation cycles

- Develop C3: Event further
- Refine other categories

Quantification

- RQ2-(How) do representations of the lone-wolf terrorist change?
 - Quantify representations, compare diachronically and interpret
 - e.g., 65% of the *Distinct* connection found (so far) are from the Pre-peak corpus – what do such findings suggest?

Reflection

Suggestions based on the challenges I've faced

1. Avoid diverging from a discourse analysis to a semantic analysis
 - Objective (research questions) is to reveal **discourse prosodies**, NOT to define finely distinguished semantic categories. Doing the latter is neither time-efficient nor improves productivity.
2. Maintain a bottom-up approach
 - Avoid applying theories top-down as this can lead to
 - “Procrustean influence” (Gabrielatos 2023) – i.e., making the data fit the theory => limiting in scope of observations
 - “Reading tea-leaves” scenario (Gabrielatos, personal communication)– i.e., seeing what you want to see => skewed results

References

- Gabrielatos, C. (2007) Selecting query terms to build a specialised corpus from a restricted-access database. *ICAME Journal* 31:5-43. https://icame.info/icame_static/ij31/ij31-page5-44.pdf
- Gabrielatos, C. (2023) The discourse presentation of autism in the UK press: A case of critical corpus lexicography. *Workshop on Corpus Linguistic Applications*, Aristotle University of Thessaloniki, 10 February 2023. <https://www.academia.edu/97513169>
- Gabrielatos, C. & Baker, P. (2008). Fleeing, sneaking, flooding: A corpus analysis of discursive constructions of refugees and asylum seekers in the UK press, 1996-2005. *Journal of English Linguistics* 36(1), 5-38.
- Gillings, M. & Hardie, A. (2022) The interpretation of topic models for scholarly analysis: An evaluation and critique of current practice. *Digital Scholarship in the Humanities*. OnlineFirst, 22 December 2022, <https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac075/6957052?login=true>
- Malone, D. (2020). Developing a complex query to build a specialised corpus: Reducing the issue of polysemous query terms. *Corpora and Discourse International Conference 2020*. <https://doi.org/10.13140/RG.2.2.31214.43846>
- Sinclair, J. (1991) *Corpus Concordance Collocation*. Oxford University Press.
- Stubbs, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.
- Van Leeuwen, T. (2008). *Discourse and Practice: New Tools for Critical Discourse Analysis*. Oxford University Press.

Contact



daniel.malone012@gmail.com



@Dan__Malone



<https://independent.academia.edu/DanielMalone14>