The Macquarie Laws of War Corpus (MQLWC): Design, Construction and Use

Annabelle Lukin¹ · Rodrigo Araujo e Castro^{1,2}

Accepted: 25 January 2022 / Published online: 19 February 2022
© Crown 2022

Abstract

This paper discusses the creation and use of the new Macquarie Laws of War Corpus (MQLWC). The corpus consists of the 110 documents of international war law stored in the International Committee of the Red Cross treaties database, starting with the 1856 Declaration Respecting Maritime Law (Paris Declaration) and ending with the most recent amendment to the Rome Statute (2019). The new MQLWC is hosted at the Sydney Corpus Lab (sydneycorpuslab.com), via its CQWeb interface, which allows for searching of frequencies, concordance lines, and collocations. The corpus can also be downloaded for offline processing in other popular concordance programs, such as #Lancsbox and AntConc. This paper introduces the corpus, describes the process of assembling the data, and explains its limitations. The paper then demonstrates some of the ways the data can be explored using the concept of 'military necessity'. The MQLWC contributes to the growing use of corpus linguistics in legal studies, and will be of particular relevance to scholars in the field of international war law.

Keywords Macquarie Laws of War Corpus · MQLWC · International war law · International humanitarian law · Military necessity

1 Introduction

This paper introduces the new Macquarie Laws of War Corpus, or MQLWC. The term 'Laws of War', also known as the 'Law of Armed Conflict' (LOAC) and 'International Humanitarian Law' (IHL) [8], refers to a body of international treaties and documents written in principle to regulate the use of force in international and non-international armed conflicts, although, as discussed below, the history and function

✉ Annabelle Lukin
annabelle.lukin@mq.edu.au

¹ Department of Linguistics, Macquarie University, Sydney, Australia

² Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

of this body of law is very much contested. The MQLWC encompasses the documents held in a treaties database by the International Committee of the Red Cross (ICRC). The ICRC maintains its IHL Treaties database as one of two sources of international war law: the second is a database of customary international law. The provision of information about legal frameworks for the regulation of international conflict is central to the ICRC's commitment to "prevent suffering by promoting and strengthening humanitarian law and universal humanitarian principles",¹ which includes a role in "reminding authorities of their legal obligations under international humanitarian law and international human rights law".² The ICRC maintains this IHL Treaties database³ as part of its mission, and this corpus consists of all text documents hosted in this database (except for one text—see below).

The Macquarie Laws of War Corpus provides new potential for conducting research into this extremely significant body of international law. While previously an "esoteric area of international law"—the preserve of the ICRC, military lawyers and a small number of academics—the field of international war law is now mainstream [22]. The factors for this transformation are argued to include changes in media reporting of war, the pervasiveness of smartphones, and the proliferation of international tribunals and courts which have "raised global expectations that those responsible for atrocities should be held accountable" [22]. Currently the texts constituting this body of law are available at the ICRC website only as a set of separate documents. They can be opened as pdf documents, or accessed on the website, where often their content is distributed across a number of links, based on text segments such as Articles of the text. This display prevents even searching for a single term across one text. These modes of access allow only for the most minimal of engagement with these significant documents. The recent addition of an 'advanced search' function to the database enables users to locate words and phrases in texts, but does not offer the advanced functionality of standard corpus linguistic tools.

With the growing interest in international war law have come debates about the history and function of the laws of war, with some scholars rejecting "the standard Western-European-centric view" characterized by the "grand narrative of international law as the purveyor of peace and civilization to the whole world" [21]. Critical accounts of international war law argue that this body of law in its various forms does not constrain the use of violence, but rather is a crucial source of its legitimation [20, 24, 26, 28]. Mégret for instance argues that treaties, statutes and customary war law sources "enable, constitute and perpetuate" war and have produced "the basic building blocks of the international grammar of violence" [26]. Jochnick and Normand [20] argue that the laws of war have in fact "facilitated rather than restrained wartime violence". At the very least, the development of this body of legal text has produced many "unintended consequences" [11], suggesting the need to study these texts as a form of public discourse beyond whatever claims might be made for their legal potential to constrain the use of geopolitical violence.

¹ See <https://www.icrc.org/en/mandate-and-mission>.

² See <https://www.icrc.org/en/what-we-do/building-respect-ihl>.

³ See <https://ihl-databases.icrc.org/ihl>.

How do one and the same texts afford such profoundly distinct interpretations? While legal analysis and debate offer one means to consider this paradox, the gaze from linguistics is indispensable for a question of this kind. As Goodrich notes, "legal discourse, like any other of the traditional rhetorical genres or language varieties, is an historically and rhetorically organised product"; as such, it can be subjected to a "critical linguistic methodology" which allows scholars to "read within the structure of legal discourse the socio-historical and political affinities and conflicts that led to the emergence of the myth of law as a unitary language and as a discrete scientific discipline" [12]. While legal texts operate in highly specialized contexts of interaction, in which those with the authority to invoke the meanings of these texts require specialized training, the texts themselves are products of the same linguistic systems on which all other interactions by speakers of the same language depend. Legal texts, as such, are subject to the same semiotic principles as any ordinary use of language. These principles include "the arbitrariness of the sign" [10], that is, the principle that linguistic signs are the unity of a concept and an acoustic image. Any potential sound can be recruited to be the expression of any meaning that humans can dream up. In addition, legal texts like any other linguistic text are the product of lexicogrammatical choices. Any word or phrase within such texts operates within multifunctional paradigms of structures which are essential to the ways these texts make meanings [18]. Finally, these texts, like all acts of meaning, have their own institutional and social-historical contexts [15]. All these features of language give it its enormous plasticity, and therefore, its ideological potential [23]. The theories and methods in linguistics which enable the study of texts of all kinds have much to offer the study of legal discourse.

The focus of this paper is to introduce the MQLWC and to demonstrate the value of corpus linguistic techniques for the study of the laws of war. Corpus linguistic methods have already begun to have a role in the study of legal discourse (see e.g. [13, 29, 30, 32]). Given the significance and indeed the growing interest in this field of international war law, the texts which are its foundation are overdue for the kind of textual study that is made possible when converted into a searchable linguistic corpus. By turning these discrete documents into a fully searchable corpus, the contest of interpretation over the efficacy of these texts will have access to new forms of evidence and argumentation made available by the standard techniques of corpus linguistics, including word frequencies, concordance patterns, and collocations. Not only does the transformation of the documents into a corpus open them up to the methods of inquiry made available by corpus linguistics, they become more readily subject to study through a variety of linguistic methods.

Corpus linguistic techniques enable researchers to efficiently search a large volume of data for a variety of linguistic patterns. With this research infrastructure, it becomes possible to determine the presence, frequency and distribution of words and phrases, as well as the typical collocation patterns of some key lexical node, within a given body of texts. In the case of the MQLWC, the corpus is not simply a representative sample of this domain: it consists of all possible texts in this

field. When available in this form, researchers can approach this corpus from both so-called corpus-driven and corpus-based or corpus-assisted methods [25]. From the perspective of corpus-driven approaches⁴—where researchers use standard techniques to search a body of texts and allow the findings to drive analysis and interpretation—it is now possible to describe a range of features of this body of law text, such as its size, and its dominant as well as its less frequent lexical patterns. From a corpus-assisted perspective, corpus methods allow a researcher to search for predefined lexicogrammatical patterns of interest to determine their frequency and distribution within international war law, with a view to testing claims about the meaningful patterns in the laws of war. Naturally, corpus-assisted methods combine well with other forms of manual text or discourse analysis which enable a deeper inquiry into linguistic patterns, though always at the cost of the scale of analysis which can be conducted. Lukin (2019) exemplifies how small scale detailed text analysis can combine with standard corpus techniques in the study of ideology, in this case ideologies around war and violence.

2 Creating the Macquarie Laws of War Corpus (MQLWC)

To construct this corpus, we first downloaded all documents archived by the ICRC from the IHL Treaties database, with the exception of a small set of texts which were directly available in plain text format via Wikisource (see “Appendix” for a full list of the documents in the corpus). All documents downloaded from the ICRC site were in pdf format, and they varied in type and in formatting. Some were not searchable. A shell script was used for converting the documents into a searchable txt format, and metadata from within the documents was removed using regular expressions in the text editor Notepad++. The files were edited where necessary to include only the text of the treaty or document: headings, tables of contents, and page numbers were removed.

The documents are organised by the ICRC by the following categories: date, topic, and by state (listing which conventions have been signed/ratified by each state). We retained the categories of date and topic. While the details of which states are parties to which convention is important, we decided that it was too difficult to implement this category. In addition, these documents are tagged by the ICRC with the following themes, which we also retained in the development of this corpus: ‘victims of armed conflict’, ‘methods and means of warfare’, ‘naval and air warfare’, ‘cultural property’, ‘criminal repression’, and ‘other treaties relating to IHL’. The database also includes a set of texts labelled ‘historical treaties and documents’: these are documents that are either no longer in force or are historically significant but which have never functioned as an international agreement, for example, the Leiber Code, described as the “earliest official government codification of the laws of war” [7]. In putting this corpus together, we decided that all these documents

⁴ Like McEnery and Hardie (McEnery and Hardie, 2012) we reject the assumptions of corpus-driven approaches, namely that the corpus itself can set the terms of the inquiry.

should be included for the role they have played in the development of accepted norms and principles of the laws of war. The Leiber Code, for instance, creates the concept of 'military necessity' [7] which remains a current principle in international war law (see below).

The corpus was then tagged to be suitable for the CQPweb interface, a software interface that underpins a number of widely used corpora, such as the British National Corpus (BNC). CQPweb (where CQP stands for Corpus Query Processor) is a 'fourth generation' corpus linguistics software system, providing usability (including for researchers with minimal or no programming skills), power and flexibility to maximise the potential of corpus data to be studied through corpus methods [19, 25]. For example, word frequencies within the corpus can be measured, both as a raw number and as a normalised frequency to enable comparison with other data sets (by default set at words per million or wpm). Concordance lines i.e. a line of text in which a keyword appears, can be retrieved, allowing for a researcher to examine words or phrases in their immediate context. In addition, the system can generate the collocates of a key word or phrase, that is, the typical words that co-occur with the word or phrase under study, with a variety of statistical measures of collocation available to the user.

To prepare the documents for upload to CQPweb, the data was lemmatized (i.e. lexical items from the same word stem were grouped together regardless of their inflectional forms) and tagged with part of speech and semantic categories using the following tagsets: CLAWS 6 part-of-speech tagger, Simple POS (using Oxford Simplified Tagset), USAS (UCREL Semantic Analysis System),⁵ and Simplified USAS. It was then uploaded into the CQPweb interface. The corpus is also available from this same location to be downloaded for offline processing by selecting "export corpus" from the sidebar menu. This allows the data to be analysed with widely used desktop programs such as #Lancsbox [5] or AntConc [2].

3 Overview of the Corpus

At the time of publication, the ICRC treaties database contained 111 documents. As one of these documents consists only of an image of an identity card to be carried by journalists (Annex II to Additional Protocol I of the Geneva Conventions), it was excluded, leaving a total of 110 documents in the corpus. The size of a corpus is typically measured through a count of tokens, although there is no standardized definition for the category of token. The variation in the protocols for counting tokens in a corpus includes decisions relating to punctuation, clitics (e.g. contractions such as "can't", or "she'll"), and multi-word expressions [6]. The token count of the MQLWC in CQPweb is nearly 392 K, as punctuation items are included as tokens. The number of types—unique words or character clusters—is 8557 (giving a type/token ratio of 0.022). When tokens are counted in a tool that ignores punctuation, such as #Lancsbox [5], the corpus is roughly 10% smaller, with a token count

⁵ Full list of categories can be found here: http://ucrel.lancs.ac.uk/usas/semtags_subcategories.txt.

of around 355 K, and a slightly reduced type count of 8513 (producing a type to token ratio of 0.024). The variation in tokenization principles across these widely used systems is a reminder that all units in linguistics are derivative.

Despite the variation in these counts, it is now possible to estimate the size of the body of documents constituting this field of law. Taking for example the #Lancsbox calculation, we can now determine that over a period of 165 years of recognised treaty-making processes, states have produced around 355,000 words purportedly with the purpose of attempting to limit the use of violence in international and non-international armed conflict.

The CQPweb user interface provides various options for searching the data: word frequencies, collocations, standard whole-of-corpus and restricted queries based on internal parameters of the corpus. In the case of MQLWC, these include the ICRC categories, and the distinction between treaties still currently in force, versus those which have lapsed or which have no legal hold. This means words or phrases can be searched across the whole data set, or in a subset of the data based on identified sub-categories such as 'victims of armed conflict', or 'criminal repression'.

As mentioned, the corpus includes all texts identified by the ICRC as falling within the scope of International Humanitarian Law. Each of these texts has played its unique role in the development of this field of law. Hence, it was decided that text IDs needed to readily identify each document, including its place in the timeframe from the first treaty document in 1856 to the most recent text at the time of writing (the 2019 amendment to the Rome Statute). Therefore, text IDs begin with the year of adoption (or where the text was never formally adopted, its year of publication), combined with the document's official name, abbreviated in some cases to give only the information necessary to identify the document. For example, the 2006 Convention for the Protection of all Persons from Enforced Disappearance has the text ID "2006_ConventionProtectionEnforcedDisappearance".

4 Limitations

The main limitation of the MQLWC corpus is that it consists only of English documents. With a few exceptions, all documents from the 1856 Paris Declaration until the 1935 Treaty on the Protection of Artistic and Scientific Institutions and Historic Monuments (Roerich Pact) were in French only, which means the documents in this corpus from this period are translations of original legal texts. Apart from the 1863 Leiber Code, English was first added as a second language in a 1922 treaty (the Washington Treaty relating to the Use of Submarines and Noxious Gases). From 1935 onwards, with occasional exceptions, English has been a recognised language for treaty documents, with Spanish (from 1935), Russian (from 1945), Chinese (from 1946) and Arabic (from 1978) gradually added as official languages.⁶ The monolingual nature of this corpus is a clear limitation. In addition, the corpus

⁶ Portuguese was an official language of a single document, the 1935 Treaty on the Protection of Artistic and Scientific Institutions and Historic Monuments (Roerich Pact).

confines itself to one source of the laws of war. There are other sources of international war law, such as army manuals and case law. Such texts are not included in the corpus, but they should be considered for a future, larger corpus of documents in the field of international war law.

5 Using the Corpus for Research

Hosted at the Sydney Corpus Lab (sydneycorpuslab.com), this dataset is now available to any researcher interested in patterns of meaning in the international laws of war and it is our hope that the corpus will enable collaborations between linguists and scholars of international war law. We offer a brief example here of how the corpus and some corpus techniques might be used to consider the argument that these laws create an entitlement to the use of violence by the states who are party to them. This has been argued for instance in Jochnick and Normand [20] who "challenge[] the notion that the laws of war serve to restrain or 'humanize' war" arguing they have instead been "formulated deliberately to privilege military necessity at the cost of humanitarian values" [20]. As a consequence, they argue, these laws facilitate rather than restrain violence, and provide it with the very valuable semiotic commodity of legitimacy [20].

Given this claim, what can this corpus reveal on the question of whether the laws of war privilege military necessity at the cost of humanitarian values? A full response to this question is beyond the scope of this paper, and requires analysis over and beyond that offered by corpus linguistics—see for instance Lukin [24] for an analysis of the lexicogrammatical structures in the definition of “war crimes” in the Rome Statute,⁷ including a discussion of the grammar that enables an unequivocal rejection of belligerent technology (the use of chemical weapons) compared with the grammar which defends the use of other lethal technologies. But here we set out some examples of how the corpus could be used to provide empirical evidence pertaining to a question of this kind. We briefly explore the presence of the concept ‘military necessity’ in the corpus, and begin to describe the semantic landscape in which this very powerful concept comes into being and through which it is sustained. A first simple step in this process is to calibrate the frequency of the term ‘military necessity’ in international war law via the MQLWC. With the new corpus available through this interface, we queried the corpus and retrieved a total of 43 instances of the term ‘military necessity’, in 20 of the 110 texts in the corpus, beginning with the 1863 Leiber Code and ending with the 2010 amendment to Article 8 of the Rome Statute.

At first glance the instances of this term do not seem particularly numerous: a frequency of 43 would give this term a rank in the late 900 s in this corpus. But a raw frequency of 43 instances in this corpus translates into a relative frequency of 109 words per million (wpm). By comparison, the frequency of ‘military necessity’ in

⁷ Rome Statute of the International Criminal Court, done at Rome 17 July 1998; in force 1 July 2002; 2187 UN Treaty Series 90.

Fig. 1 The concept of 'military necessity' in the google books corpus

the BNC is 0.07 wpm (and the concordance lines show fewer than half of the usages in the BNC refer to 'military necessity' as a principle). Google's Ngram Viewer [27] shows the peaks of the use of this concept in its books corpus as 1864, 1917, and 1944 (see Fig. 1)—the American Civil War and the First and Second World Wars. The highest peak is 1944, where the figure is 0.0007165 per hundred (the standard reporting metric of the Google Books Ngram Viewer), or 7 wpm. The News of the Web corpus [9], a corpus of English news text collected from 1/1/2010 until the present from 20 countries, contains 451 instances across its 13.8 billion words, which translates into a normalised frequency of around 0.0033 words per million (search conducted on 23/11/21). In the context of international war law, the mere 43 instances constitute a very high frequency for this concept.

It perhaps seems an obvious finding that the concept of 'military necessity' would be more frequent in this corpus than in other comparison corpora. But what this finding confirms is that the term is register-specific [3, 17]. It was born out of, and resides within, highly specialized and highly abstract discourse. That 'military necessity' lives in this register confirms its character as an abstract legal principle, one that requires the highly elaborated semiotic edifice of legal discourse to survive. When its legal role is to defend actions that, without the cover of this principle could be classified as 'war crimes', it is almost breathtaking to understand the fundamentally derivative character on this concept. It is, as Halliday describes, a "construct of virtual reality", that is, of "a semiotic alternative universe" [16] being crucially at a distance from the material events to which it can become attached. An observer 'on the ground' cannot, with the naked eye, determine whether acts of violence are 'war crimes' or are permissible under 'military necessity'—only legal argumentation can make such determinations.

To treat the texts on international war law as a linguistic corpus arguably enables researchers and practitioners to 'deautomatize' [31] these weighty legal principles. For example, 'military necessity' is a simple collocation of two terms. 'Necessity' is a noun, and 'military' functions as classifying adjective

Table 1 Twenty most frequent open system items in the MQWLC

Rank	Lexical item	Raw frequency	Normalised frequency	Text dispersion (n/total n)
1	Convention	1719	4387 wpm	77/110
2	Parties	1645	4198	85/110
3	State	1552	3961	68/110
4	Article	1443	3683	80/110
5	WAR	1397	3565	79/110
6	Party	1351	3448	63/110
7	International	1154	2945	90/110
8	States	1049	2677	65/100
9	Power	975	2488	52/110
10	Present	846	2159	78/110
11	Prisoners	832	2123	34/110
12	MILITARY	827	2111	66/110
13	Conflict	791	2019	42/110
14	COURT	759	1937	37/110
15	Persons	716	1827	67/110
16	Paragraph	675	1723	68/110
17	PROVISIONS	674	1720	78/110
18	Powers	665	1697	57/110
19	Contracting	664	1695	60/110
20	Armed	643	1641	62/110

premodifying 'necessity' [14]. Thus, the conjunction of these lexical items creates a subtype of 'necessity', that of the 'military' kind. By such a simple grammatical act, a powerful binding legal concept was manufactured, but only through the force of the kind of texts in which it has been created and reiterated.

Given the importance of the semantic environment of this concept, other insistent features of this data are likely to assist in understanding the power and plausibility of this principle. Table 1 sets out the top 20 most frequent lexical items in the corpus (with closed system grammatical items filtered out). Many of the higher-ranked items in this list are references to key participants in the treaties (Parties, Party, state/s), or to the documents or aspects of the document (Convention, article, and 'present', which almost entirely appears for deictic purposes, e.g. "the *present* convention/article/statute/protocol"). If we set aside references to the key participants and documents, the word 'military' now sits in the top five lexical items of the corpus. Its normalised frequency is 2111 wpm. For the sake of comparison, in the BNC 'military' has a frequency of 114 wpm: the term 'military' is 18 times more frequent in the MQLWC. This corpus is, therefore, a key semiotic environment for creating or reinforcing the meanings of 'military',

Table 2 Top 20 collocates of 'military' in MQLWC

Rank	Collocate	Total no	Expected collocate frequency	Observed collocate frequency	In no. of texts	MI3 value
1	Objectives	106	2.244	84	17	18.011
2	Operations	199	4.213	82	24	16.998
3	Necessity	88	1.863	43	19	15.381
4	Objective	55	1.164	36	10	15.29
5	Aircraft	431	9.125	69	6	15.136
6	Advantage	42	0.889	29	12	14.743
7	Authorities	233	4.933	47	15	14.361
8	Action	100	2.117	27	12	13.183
9	Imperative	27	0.572	17	9	13.069
10	Authority	176	3.726	31	14	12.965
11	Permit	86	1.821	24	14	12.89
12	Direct	103	2.181	25	11	12.807
13	Considerations	22	0.466	14	9	12.524
14	Military	826	17.488	46	14	12.442
15	Purposes	207	4.383	29	18	12.442
16	Concrete	12	0.254	11	7	12.355
17	Contribution	28	0.593	14	7	12.176
18	Anticipated	14	0.296	11	7	12.133
19	Hospital	178	3.769	25	9	12.018
20	Auxiliary	39	0.826	15	1	11.997

including its collocation with 'necessity', providing further evidence for the register-specific nature of this concept.

Table 2 shows the top 20 collocates of 'military'. Following Brezina et al. [4], the collocation parameter notation for this set of collocates is: MI3(9), R5-L5, C5-NC5, closed system items removed (where 'closed system' means those belonging to grammatical paradigms, see Halliday and Matthiessen [14]. In other words, the collocates are based on a window of five words to the left and right of 'military', with a minimum requirement of five instances of each collocate and of their collocation with 'military', adopting the collocation measure Mutual Information 3, which combines both mutuality and frequency [4]. Following these parameters, 'necessity' ranks as the third highest collocate of 'military'. Of the 88 instances of 'necessity', very nearly 50% combine with 'military'. From the perspective of 'necessity', 'military' is the number 1 collocate. Thus it is not simply that 'military' attracts 'necessity', but that in this context, 'necessity' attracts the word 'military'. Other collocates of 'military' reveal close synonyms of 'military necessity', such as 'advantage', 'considerations' and 'imperative'. In addition, the collocates show a deep legitimating semantics around the concept of 'military', including with lexical items like 'objectives', 'operations', 'objective', 'authorities', 'authority', 'permit', and 'purposes'. The environment for the most

5	1863_LeiberCode	moral beings , responsible to one another and to God . .	Military necessity	does nc
6	1863_LeiberCode	but disclaims acts of perfidy ; and , in general ,	military necessity	does nc
7	1863_LeiberCode	offenses of the owner , can be seized only by way of	military necessity	, for th
8	1868_Geneva_AdditionalArticlesRelatingCo	only be allowed to delay for a short time in case of	military necessity	. . Arra
9	1880_Oxford_LawsWarLand_Oxford	however he can only delay for a short time in case of	military necessity	. . Mea
10	1880_Oxford_LawsWarLand_Oxford	detained for such length of time only as is warranted by strict	military necessity	. (f) S
11	1880_Oxford_LawsWarLand_Oxford	occupant . Their destruction is forbidden , unless it be demanded by	military necessity	. They ;
12	1880_Oxford_LawsWarLand_Oxford	or science , is formally forbidden , save when urgently demanded by	military necessity	. (b) P
13	1906_Geneva_AmeliorationConditionWounded	within such period and by such route as may accord with	military necessity	. They ;
14	1906_Geneva_AmeliorationConditionWounded	operations , however , may use them , in case of important	military necessity	, if , be
15	1906_Geneva_AmeliorationConditionWounded	1 . A belligerent intercepting a convoy may , if required by	military necessity	, break
16	1909_London_LawsNavalWar	and wounded . They can , however , in case of urgent	military necessity	and sut
17	1923_Hague_WirelessTelegraphWarAirfare	private aircraft may not be destroyed except in the event of a	military necessity	of extre
18	1929_Geneva_WoundedSickArmiesField	the field may make use of them , in case of urgent	military necessity	, provic
19	1945_London_CharterNuremberg	of cities , towns or villages , or devastation not justified by	military necessity	; (c) ')
20	1949_1_First_Geneva_Convention_WoundedSi	the field may make use of them , in case of urgent	military necessity	, provic
21	1949_1_First_Geneva_Convention_WoundedSi	and extensive destruction and appropriation of property , not justified by	military necessity	and car
22	1949_2_Second_Geneva_Convention_WoundedS	accommodated therein , apply them to other purposes in case of urgent	military necessity	. Any h
23	1949_2_Second_Geneva_Convention_WoundedS	and extensive destruction and appropriation of property , not justified by	military necessity	and car
24	1949_3_Third_Geneva_Convention_POWs	restricted . Visits may not be prohibited except for reasons of imperative	military necessity	, and th
25	1949_4_Fourth_Geneva_Convention_Civilian	obligations imposed upon it by virtue of the present Convention . Should	military necessity	require
26	1949_4_Fourth_Geneva_Convention_Civilian	Such visits may not be prohibited except for reasons of imperative	military necessity	, and th
27	1949_4_Fourth_Geneva_Convention_Civilian	hostages and extensive destruction and appropriation of property , not justified by	military necessity	and car

Fig. 2 Sample concordance lines for 'military necessity' showing text locations in MQLWC

Table 3 Texts in which the term "military necessity" appears in the MQLWC

File	Tokens	Frequency
1863_LeiberCode.txt	9277	7
1868_Geneva_AdditionalArticlesRelatingConditionWoundedin War.txt	1437	1
1880_Oxford_LawsWarLand_Oxford.txt	4932	4
1906_Geneva_AmeliorationConditionWoundedSickArmiesField.txt	2965	3
1909_London_LawsNavalWar.txt	5109	1
1923_Hague_WirelessTelegraphWarAirfare.txt	5195	1
1929_Geneva_WoundedSickArmiesField.txt	3664	1
1945_London_CharterNuremberg.txt	3223	1
1949_1_First Geneva Convention_WoundedSickArmedForcesField.txt	8609	2
1949_2_Second Geneva Convention_WoundedSickSea.txt	6771	2
1949_3_Third Geneva Convention_POWs.txt	20,189	1
1949_4_Fourth Geneva Convention_Civilians.txt	21,367	3
1950_PrinciplesILNuremberg.txt	386	1
1954_Hague_Conv__ProtectionCulturalProperty.txt	7899	2
1977_Geneva_Convention_Additional Protocol I.txt	21,806	4
1993_StatuteCriminalTribFormerYugoslavia.txt	3774	2
1994_SamRemo_Manual_ArmedConflictsatSea.txt	10,661	2
1998_Rome Statute of the International Criminal Court.txt	31,001	1
1999_Hague_SecondProtocolCulturalProperty.txt	6137	3
2010_RomeStatute_Art8Amendment.txt	1667	1

Solution 1 to 43 Page 1 / 1	
24 April 1863 : Martial law -- Military jurisdiction --	Military necessity -- Retaliation A place , district , or country occupied by an
well as in the dictation of general laws , as far as	military necessity requires this suspension , substitution , or dictation . The commander of
by statute on courts-martial , are tried by military commissions . .	Military necessity , as understood by modern civilized nations , consists in the necessity
lawful according to the modern law and usages of war . .	Military necessity admits of all direct destruction of life or limb of " armed
moral beings , responsible to one another and to God . .	Military necessity does not admit of cruelty -- that is , the infliction of
but disclaims acts of perfidy ; and , in general ,	military necessity does not include any act of hostility which makes the return to
offenses of the owner , can be seized only by way of	military necessity , for the support or other benefit of the army or of

Fig. 3 Concordance lines for 'military necessity' in the leiber code

powerful conjunction of 'military' and 'necessity' is also one in which the military is constructed as organised, authoritative, and purposeful.

Finally, we comment on the potential of concordance lines as a means to more deeply explore the use and dispersion of this term across the corpus. Concordance lines can show us which of the 110 documents use this term (see example concordance lines in Fig. 2; and see Table 3 for a summary of the dispersion of the term across 20 texts in the corpus). This dispersion reveals the longevity of the concept, beginning in 1863 and still present in a 2010 amendment to the Rome Statute: it has lasted nearly 150 years in the laws of war. The 1863 Leiber Code officially launches the concept and so perhaps unsurprisingly has the most number of instances, seven in total (or 16% of the total instances). By reviewing all instances of the term in the corpus, and in particular by reviewing the window to the right of the term, we see that only in the Leiber Code is the term subject to some kind of definition or explanation, including that it "consists in the necessity of ...etc. " (3rd line), "admits of all direct destruction of life" (4th line), "does not admit of cruelty" (5th line), "does not include any act of hostility" (6th line). (Fig. 3) No other instance of the term in this corpus includes any further attempt at definition or explanation. It is remarkable that in over 110 documents of international war law, and across some 150 years, this body of texts displays so little interest in the meaning or limits of this profoundly powerful concept.

In addition to the Leiber Code, the term turns up in all four of the 1949 Geneva Conventions, as well as the 1977 Additional Protocol I, the two documents pertaining to the Nuremberg trials, and the statute establishing the tribunal to try war crimes committed in the former Yugoslavia. In addition, one instance is found in a 2010 amendment to the Rome Statute. It also appears in documents such as the 1880 Oxford Laws of War on Language and the 1909 London Declaration on the Laws of Naval War. In terms of the distinction between so-called "Hague law" (law regulating how war/violence is used by states) and "Geneva law" (law regulating the treatment of victims of war) (see e.g. [8]), the concept moves easily between both fields of war law. Perhaps also noteworthy is that another significant document for this concept is the 1977 Additional Protocol 1, with its focus on "provisions protecting the victims of armed conflicts". In a paper arguing against US ratification of this protocol, Major Guy Roberts claims the 1977 Additional Protocol 1 was "an attempt to shift the balance established between military necessity and humanitarian principles in such a way as to hamper the ability of states to use military force to attain political objectives" (cited in [1])—yet nearly 10% of all references to this concept

across 150 years of international war law documents turn up in Additional Protocol I.

Above we have presented a few simple analyses that are easily generated once the relevant texts have been formed into a corpus, and some software system is available for their processing. These analyses show the strong presence of 'military' in this corpus, and the mutual attraction between 'military' and 'necessity' here. By comparing the frequency of 'military necessity' in this specialized corpus with a variety of other corpora (the BNC, the Google Books Corpus, and the News on the Web corpus), we have shown the register-dependent nature of this concept—it was created and has been perpetuated largely through this body of international law, suggesting that international war law plays a fundamental role in legalizing geopolitical violence. In addition, we have reported the longevity of the concept across nearly 150 years of international law and shown the various related collocations that all work to legitimize the place of the military in this discourse, as well as to reiterate the principle of 'military necessity' through other wording. We anticipate the corpus will be a basis for deeper linguistic work on this field of law and will create improved opportunities for interdisciplinary work on the role of legal discourse in relation to war and violence.

6 Conclusion

This paper has introduced the MQLWC and explored how simple corpus linguistic methods can reveal important dimensions of meaning in this data. These methods have been used within a corpus-assisted approach—that is, they have to be used to explore the data with a question already in mind, namely, the place of 'military necessity' in this body of law. The question was derived from critical analyses of standard claims about the rise and function of international war law [20, 24, 26, 28]. This is one of many questions which could be brought to this data, including, for instance, how the texts of international war law construct the category 'civilian', how gender operates within international war law, or how some kinds of beligerent technologies are proscribed, while others are allowed explicitly or implicitly to become a naturalized part of geopolitical violence. Even with these simple techniques, important semiotic patterns in this data can be elucidated. Corpus methods alone have their limitations—word frequencies and collocations can appear unmotivated or disconnected from the institutional work that texts of this kind perform. But when combined with an understanding of the political and historical context of these texts, and with a linguistic account of the role of language in legal contexts, corpus tools and methods are invaluable. The potential to study these significant legal documents with the power of linguistics, including corpus linguistics, is well overdue.

Appendix

See Table 4.

Table 4 List of documents in the corpus

Date of adoption	Full title
1856	Declaration Respecting Maritime Law, Paris, 16 April 1856
1863	Instructions for the Government of Armies of the United States in the Field (Leiber Code) 24 April 1863
1863	Resolutions of the Geneva International Conference. Geneva, 26–29 October 1863
1864	Convention for the Amelioration of the Condition of the Wounded in Armies in the Field. Geneva, 22 August 1864
1868	Additional Articles relating to the Condition of the Wounded in War. Geneva, 20 October 1868
1868	Declaration Renouncing the Use, in Time of War, of Explosive Projectiles Under 400 Grammes Weight, St Petersburg 29 November/ 11 December 1868
1874	Project of an International Declaration concerning the Laws and Customs of War. Brussels, 27 August 1874
1880	The Laws of War on Land. Oxford 9 September 1880
1899	Final Act of the International Peace Conference, The Hague 29 July 1899
1899	Convention (II) with Respect to the Laws and Customs of War on Land and its annex: Regulations concerning the Laws and Customs of War on Land. The Hague, 29 July 1899
1899	Declaration (IV,1), to Prohibit, for the Term of Five Years, the Launching of Projectiles and Explosives from Balloons, and Other Methods of Similar Nature. The Hague, 29 July 1899
1899	Declaration (IV,2) concerning Asphyxiating Gases. The Hague, 29 July 1899
1899	Declaration (IV,3) concerning Expanding Bullets. The Hague, 29 July 1899
1899	Convention (III) for the Adaptation to Maritime Warfare of the Principles of the Geneva Convention of 22 August 1864. The Hague, 29 July 1899
1904	Hague Convention on Hospital Ships. The Hague, 21 December 1904
1906	Convention for the Amelioration of the Condition of the Wounded and Sick in Armies in the Field. Geneva, 6 July 1906
1907	Convention (IV) respecting the Laws and Customs of War on Land and its annex: Regulations concerning the Laws and Customs of War on Land. The Hague, 18 October 1907
1907	Convention (IX) concerning Bombardment by Naval Forces in Time of War. The Hague, 18 October 1907
1907	Declaration (XIV) Prohibiting the Discharge of Projectiles and Explosives from Balloons. The Hague, 18 October 1907
1907	Final Act of the Second Peace Conference, the Hague, 18 October 1907
1907	Convention (X) for the Adaptation to Maritime Warfare of the Principles of the Geneva Convention. The Hague, 18 October 1907
1907	Convention (XII) relative to the Creation of an International Prize Court. The Hague, 18 October 1907
1907	Convention (VI) relating to the Status of Enemy Merchant Ships at the Outbreak of Hostilities. The Hague, 18 October 1907
1907	Convention (VII) relating to the Conversion of Merchant Ships into War-Ships. The Hague, 18 October 1907
1907	Convention (VIII) relative to the Laying of Automatic Submarine Contact Mines. The Hague, 18 October 1907
1907	Convention (XI) relative to certain Restrictions with regard to the Exercise of the Right of Capture in Naval War. The Hague, 18 October 1907

Table 4 (continued)

Date of adoption	Full title
1907	Convention (XIII) concerning the Rights and Duties of Neutral Powers in Naval War. The Hague, 18 October 1907
1907	Convention (III) relative to the Opening of Hostilities. The Hague, 18 October 1907
1907	Convention (V) respecting the Rights and Duties of Neutral Powers and Persons in Case of War on Land. The Hague, 18 October 1907
1909	Final Protocol to the Naval Conference of London, 26 February 1909
1909	Declaration concerning the Laws of Naval War. London, 26 February 1909
1910	Additional Protocol to the Convention relative to the Establishment of an International Prize Court. The Hague, 19 September 1910
1913	Manual of the Laws of Naval War. Oxford, 9 August 1913
1922	Treaty relating to the Use of Submarines and Noxious Gases in Warfare. Washington, 6 February 1922
1923	Rules concerning the Control of Wireless Telegraphy in Time of War and Air Warfare. Drafted by a Commission of Jurists at the Hague, December 1922—February 1923
1925	Protocol for the Prohibition of the Use of Asphyxiating, Poisonous or Other Gases and of Bacteriological Methods of Warfare. Geneva, 17 June 1925
1928	Convention on Maritime Neutrality. Havana, 20 February 1928
1929	Final Act of the Diplomatic Conference. Geneva, 27 July 1929
1929	Convention for the Amelioration of the Condition of the Wounded and Sick in Armies in the Field. Geneva, 27 July 1929
1929	Convention relative to the Treatment of Prisoners of War. Geneva, 27 July 1929
1930	Treaty for the Limitation and Reduction of Naval Armaments, (Part IV, Art. 22, relating to submarine warfare). London, 22 April 1930
1934	First draft Convention adopted in Monaco (Sanitary cities and localities), 27 July 1934
1934	Draft International Convention on the Condition and Protection of Civilians of enemy nationality who are on territory belonging to or occupied by a belligerent. Tokyo 1934
1935	Treaty on the Protection of Artistic and Scientific Institutions and Historic Monuments (Roerich Pact). Washington, 15 April 1935
1936	Procès-verbal relating to the Rules of Submarine Warfare set forth in Part IV of the Treaty of London of 22 April 1930. London, 6 November 1936
1937	The Nyon Agreement, 14 September 1937
1937	Agreement supplementary to The Nyon Agreement. Geneva 17 September 1937
1938	Draft Convention for the Protection of Civilian Populations Against New Engines of War, Amsterdam, 1938
1945	Agreement for the Prosecution and Punishment of the Major War Criminals of the European Axis, and Charter of the International Military Tribunal. London, 8 August 1945
1946	Affirmation of the Principles of International Law recognised by the Charter of the Nüremberg Tribunal. Resolution 95 (I) of the United Nations General Assembly, 11 December 1946
1948	Convention on the Prevention and Punishment of the Crime of Genocide, 9 December 1948
1949	Geneva_Final Act of the Diplomatic Conference of Geneva, 12 August 1949
1949	Convention (I) for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field. Geneva, 12 August 1949

Table 4 (continued)

Date of adoption	Full title
1949	Convention (II) for the Amelioration of the Condition of Wounded, Sick in Shipwrecked Members of Armed Forces at Sea. Geneva, 12 August 1949
1949	Convention (III) relative to the Treatment of Prisoners of War. Geneva, 12 August 1949
1949	Convention (IV) relative to the Protection of Civilian Persons in Time of War. Geneva, 12 August 1949
1949	Resolutions of the Diplomatic Conference. Geneva, 12 August 1949
1950	Principles of International Law Recognized in the Charter of the Nuremberg Tribunal and in the Judgment of the Tribunal, 1950
1954	Final Act of the Intergovernmental Conference on the Protection of Cultural Property in the Event of Armed Conflict. The Hague, 14 May 1954
1954	Convention for the Protection of Cultural Property in the Event of Armed Conflict. The Hague, 14 May 1954
1954	Protocol for the Protection of Cultural Property in the Event of Armed Conflict. The Hague, 14 May 1954
1954	Resolutions of the Intergovernmental Conference on the Protection of Cultural Property in the Event of Armed Conflict. The Hague, 14 May 1954
1956	Draft Rules for the Limitation of the Dangers incurred by the Civilian Population in Time of War. ICRC, 1956
1968	Convention on the Non-Applicability of Statutory Limitations to War Crimes and Crimes Against Humanity, 26 November 1968
1968	Human Rights in Armed Conflict. Resolution XXIII adopted by the International Conference on Human Rights. Tehran, 12 May 1968
1968	Respect for Human Rights in Armed Conflicts. Resolution 2444 (XXIII) of the United Nations General Assembly, 19 December 1968
1969	The Distinction between Military and Non-Military Objectives in General and Particularly the Problems Associated with Weapons of Mass Destruction. Institute of International, Edinburgh, 9 September 1969
1971	Convention on the Prohibition of the Development, Production and Stockpiling of the Bacteriological (Biological) and Toxin Weapons and on their Destruction. Open for Signature at London, Moscow and Washington. 10 April 1972
1974	European Convention on the Non-Applicability of Statutory Limitations to Crimes against Humanity and War Crimes. Strasbourg, 25 January 1974
1976	Convention on the prohibition of military or any hostile use of environmental modification techniques, 10 December 1976
1977	OAU Convention for the Elimination of Mercenarism in Africa. Libreville, 3rd July 1977
1977	Final Act of the Diplomatic Conference of Geneva of 1974–1977
1977	Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977
1977	Annex I (To Protocol Additional I to the Geneva Conventions of 1949): Regulations concerning identification, 6 June 1977
1977	Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II), 8 June 1977
1977	Resolutions of the Diplomatic Conference of Geneva of 1974–1977
1979	Resolution on Small-Calibre Weapons Systems. Geneva, 28 September 1979

Table 4 (continued)

Date of adoption	Full title
1980	Final Act of the United Nations Conference on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May be Deemed to be Excessively Injurious or to Have Indiscriminate Effects. Geneva, 10 October 1980
1980	Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May be Deemed to be Excessively Injurious or to Have Indiscriminate Effects. Geneva, 10 October 1980
1980	Protocol on Non-Detectable Fragments (Protocol I). Geneva, 10 October 1980
1980	Protocol (II) on Prohibitions or Restrictions on the Use of Mines, Booby-Traps and Other Devices. Geneva, 10 October 1980
1980	Protocol on Prohibitions or Restrictions on the Use of Incendiary Weapons (Protocol III). Geneva, 10 October 1980
1989	International Convention against the Recruitment, Use, Financing and Training of Mercenaries, 4 December 1989
1989	Convention on the Rights of the Child, 20 November 1989
1993	Statute of the International Criminal Tribunal for the Former Yugoslavia, 25 May 1993
1993	Convention on the prohibition of the development, production, stockpiling and use of chemical weapons and on their destruction, Paris 13 January 1993
1993	Annex (I) to Protocol Additional I to the Geneva Conventions of 1949: Regulations concerning identification, as amended on 30 November 1993
1994	Statute of the International Criminal Tribunal for Rwanda, 8 November 1994
1994	San Remo Manual on International Law Applicable to Armed Conflicts at Sea, 12 June 1994
1995	Protocol on Blinding Laser Weapons (Protocol IV to the 1980 Convention), 13 October 1995
1996	Protocol on Prohibitions or Restrictions on the Use of Mines, Booby-Traps and Other Devices as amended on 3 May 1996 (Protocol II to the 1980 CCW Convention as amended on 3 May 1996)
1997	Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on their Destruction, 18 September 1997
1998	Rome Statute of the International Criminal Court, 17 July 1998
1999	Second Protocol to the Hague Convention of 1954 for the Protection of Cultural Property in the Event of Armed Conflict The Hague, 26 March 1999
2000	Optional Protocol to the Convention of the Rights of the Child on the involvement of children in armed conflicts, 25 May 2000
2001	Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May be Deemed to be Excessively Injurious or to Have Indiscriminate Effects. Geneva, 10 October 1980
2002	Agreement between the United Nations and the Government of Sierra Leone and Statute of the Special Court for Sierra Leone, 16 January 2002
2003	Protocol on Explosive Remnants of War (Protocol V to the 1980 CCW Convention), 28 November 2003
2005	Protocol additional to the Geneva Conventions of 12 August 1949, and relating to the Adoption of an Additional Distinctive Emblem (Protocol III), 8 December 2005
2006	Convention for the Protection of all Persons from Enforced Disappearance, 20 December 2006
2008	Convention on Cluster Munitions, 30 May 2008

Table 4 (continued)

Date of adoption	Full title
2010	Amendment to the Rome Statute of the International Criminal Court on War Crimes, amended article 8, 10 June 2010
2010	Amendment to the Rome Statute of the International Criminal Court on the Crime of Aggression, articles 8bis, 15bis and 15ter, 11 June 2010
2013	Arms Trade Treaty, 2 April 2013
2015	Amendment to article 124 of the Rome Statute of the International Criminal Court
2017	Treaty on the Prohibition of Nuclear Weapons, 7 July 2017
2017	Amendment to the Statute of the International Criminal Court, amended article 8 (Weapons which use microbial or other biological agents, or toxins)
2017	Amendment to the Statute of the International Criminal Court, amended article 8 (Weapons the primary effect of which is to injure by fragments undetectably by x-rays in the human body)
2017	Amendment to the Statute of the International Criminal Court, amended article 8 (Blinding laser weapons)
2019	Amendment to the Statute of the International Criminal Court, amended article 8 (Internationally using starvation of civilians)

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Declarations

Conflict of interest The authors have no interests financial or otherwise in regards to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Alexander, A. 2015. A short history of international humanitarian law. *European Journal of International Law* 26 (1): 109–138. <https://doi.org/10.1093/ejil/chv002>.
2. Anthony, L. 2020. AntConc (Version 3.5.9) [Computer Software]. Retrieved 9 Nov 2021. <https://www.laurenceanthony.net/software>.
3. Biber, D., and S. Conrad. 2009. *Register, genre and style*. Cambridge: Cambridge University Press.
4. Brezina, V., T. McEnery, and S. Wattam. 2015. Collocations in context: A new perspective on collocation networks. *International Journal of Corpus Linguistics* 20 (2): 139–173. <https://doi.org/10.1075/ijcl.20.2.01bre>.

5. Brezina, V., P. Weill-Tessier, and A. McEnery. 2020. #LancsBox v. 5.x. [software]. Retrieved 9 Nov 2021. <http://corpora.lancs.ac.uk/lancsbox>.
6. Brezina, V., and M. Timperley. 2017. How large is the BNC? A proposal for standardised tokenisation and word counting. Paper presented at International Corpus Linguistics Conference. University of Birmingham.
7. Carnahan, B.M. 1998. Lincoln, Lieber and the laws of war: The origins and limits of the principle of military necessity. *The American Journal of International Law* 92 (2): 213–231. <https://doi.org/10.2307/2998030>.
8. Crawford, E., and A. Pert. 2015. *International humanitarian law*. Cambridge: Cambridge University Press.
9. Davies, M. 2016. Corpus of News on the Web (NOW). <https://www.english-corpora.org/now/>. Retrieved 22 Nov 2021.
10. de Saussure, F. 1974. *Course in general linguistics* (trans: Baskin, W.). London: Fontana/Collins.
11. Fazal, T.M. 2018. *Wars of law: Unintended consequences in the regulation of armed conflict*. Ithaca and London: Cornell University Press.
12. Goodrich, P. 1990. *Legal discourse: Studies in linguistics, rhetoric and legal analysis*. London: Palgrave Macmillan.
13. Goźdz-Roszkowski, S. 2021. Corpus linguistics in legal discourse. *International Journal for the Semiotics of Law* 34: 1515–1540. <https://doi.org/10.1007/s11196-021-09860-8>.
14. Halliday, M.A.K., and C.M.I.M. Matthiessen. 2014. *An introduction to Halliday's functional grammar: fourth edition*, 3rd ed. London: Arnold.
15. Halliday, M. A. K. 1975. Language as social semiotic: Towards a general sociolinguistic theory. Paper presented at International Corpus Linguistics Conference. A. Makkai & V. Becker Makkai. Columbia, SC: Hornbeam Press.
16. Halliday, M.A.K. 2013. Why do we need to understand about language? In *Halliday in the 21st century*, ed. J.J. Webster, 71–81. London: Bloomsbury.
17. Halliday, M.A.K., and R. Hasan. 1985. *Language, context and text: Aspects of language in a social semiotic perspective*. Geelong: Deakin University Press.
18. Halliday, M.A.K., and C.M.I.M. Matthiessen. 2014. *An introduction to functional grammar: Fourth edition*, 3rd ed. London: Arnold.
19. Hardie, A. 2012. CQPweb—combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics* 17 (3): 380–409. <https://doi.org/10.1075/ijcl.17.3.04har>.
20. Jochnick, C.A., and R. Normand. 1994. The legitimization of violence: A critical history of the laws of war. *Harvard International Law Journal* 35 (1): 49–95.
21. Jouannet, E.T., and A. Peters. 2014. The journal of the history of international law: A forum for new research. *Journal of the History of International Law* 16: 1–8. <https://doi.org/10.1163/15718050-12340017>.
22. Liivoja, R., and T. McCormack, eds. 2016. *Routledge handbook of law of armed conflict*. London: Routledge.
23. Lukin, A. 2019. *War and its ideologies: A social-semiotic theory and description*. Singapore: Springer.
24. Lukin, A. 2020. How international law makes violence legal: A case study of the Rome Statute. *Language, Context and Text: The Social Semiotics Forum* 2 (1): 91–120. <https://doi.org/10.1075/langct.00022.luk>.
25. McEnery, A., and A. Hardie. 2012. *Corpus linguistics: Method, theory, practice*. Cambridge: Cambridge University Press.
26. Mégret, F. 2016. Theorizing the laws of war. In *Oxford handbook on the theory of international law*, ed. A. Orford and F. Hoffman. Oxford: Oxford University Press. <https://doi.org/10.1093/law/9780198701958.003.0038>.
27. Michel, J.-B., Y.K. Shen, A.P. Aiden, A. Veres, M.K. Gray, The Google Books Team, J.P. Pickett, D. Hoiberg, D. Clancy, P. Norvig, J. Orwant, S. Pinker, M.A. Nowak, and E.L. Aiden. 2011. Quantitative analysis of culture using millions of digitized books. *Science* 331: 176–182. <https://doi.org/10.1126/science.1199644>.
28. Normand, R., and C.A. Jochnick. 1994. The legitimization of violence: A critical history of the laws of war. *Harvard International Law Journal* 35 (2): 387–416.
29. Phillips, J.C., and J. Egbert. 2017. Advancing law and corpus linguistics: Importing principles and practices from survey and context analysis methodologies to improve corpus design and analysis. *BYU Law Review* 6: 1589–1620.

30. Potts, A., and A.L. Kjær. 2016. Constructing achievement in the international criminal tribunal for the former Yugoslavia (ICTY): A corpus-based critical discourse analysis. *International Journal for the Semiotics of Law* 29: 525–555. <https://doi.org/10.1007/s11196-015-9440-y>.
31. Shklovsky, V. 1965. Art as technique. In *Russian formalist criticism: Four essays*, ed. L.T. Lemon and M.J. Reis, 3–24. Lincoln: University of Nebraska Press.
32. Solan, L.M. 2020. Corpus linguistics as method of legal interpretation. *International Journal for the Semiotics of Law - Revue internationale de Sémiotique juridique* 33: 283–298.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.