The International Laws of War: Linguistic Analysis from the Perspectives of Register, Corpus and Grammatical Patterning

Annabelle Lukin | ORCID: 0000-0003-1023-7207
Department of Linguistics, Macquarie University, Sydney, NSW, Australia
Annabelle.lukin@mq.edu.au

Alexandra García Marrugo | ORCID: 0000-0002-2967-2738
Learning Hub Lead, University of Sydney, Sydney, NSW, Australia
Corresponding author
alexandra.garcia@sydney.edu.au

Abstract

All texts, including all legal texts, are constructed in language. All legal constructs and discriminations are an effect of language. As such, linguistics and text analysis should be considered a necessary adjunct to legal studies and, in particular, to critical legal studies. While the disciplines of linguistics and law are increasingly interacting, there is a paucity of linguistic analysis in the field of the international laws of war. This paper seeks to open doors to collaboration, by viewing the texts of the international laws of war from three linguistic perspectives: as a ‘register’, as a ‘corpus’, and in terms of their lexicogrammatical patterning. In terms of register, the international laws of war herald a new form of textual practice, the function and effects of which are subject to debate. As a corpus, some dominant lexical habits of these texts are explored and interpreted for their ideological implications and reactances. Finally, an examination of covert lexicogrammatical meanings in this register, via a small extract from Article 8 of the Rome Statute, illuminates the contradictory meanings that these texts navigate. These three perspectives offer a preliminary glimpse into the value of linguistic analysis for critical perspectives on the international laws of war.
1 Introduction

Legal texts are, by design, complex texts that require specialised training to interpret. They promulgate crucial legal concepts that are dressed in legal regalia. On the surface, these texts appear able to rise above the hurly burly of ordinary, nonlegal uses of language. As such, claims have been made for the special nature of legal discourse. Goodrich in his 1987 book *Legal Discourse* describes it as a ‘closed code’ and argues that ‘the rules of the legal grammatical code’ ‘transform’ or ‘translate’ ordinary language and meanings into this closed system. A decade earlier, the French linguist Greimas had described legal discourse as a ‘secondary metasemiotic’, with its own lexicon and syntax and a capacity to create ‘autonomous semiotic objects endowed with personality and quasi-organic functions’.

However, while legal texts have unique features, they are in most respects like any other use of language. Like any other text, a legal text depends on the semiotic principles that enable every act of meaning. They are structured and coherent acts of meaning, responsive to and constrained by the nature of their social and cultural context. Those who create, interpret, and study legal texts stand to benefit from linguistic insights into the internal semiotic machinery of language, upon which all texts depend.

This paper brings a linguistic focus to texts variously referred to as the ‘laws of war’, the ‘laws of armed conflict’ or ‘international humanitarian law’.¹ Until recently, histories of this body of international law were ‘characterized by the grand narrative of international law as the purveyor of peace and civilisation to the whole world’,² as a ‘broadly humanist and civilizing force, standing back from war, judging it to be just or unjust, while offering itself as a code of conduct to limit violence on the battlefield’.³ More recently, this ‘standard Western-European-centred view’ has been challenged, so that ‘virtually no one today thinks that he or she could produce a new grand narrative of

---

international law which would be objective, entirely neutral, absolute and universal'. Alternatives to the ‘grand narrative’ of international law as a purveyor of peace include claims that the international laws of war, rather than constraining violence and being a force for ‘civilisation’, are a potent source for the legitimation of violence.

For example, Jochnick and Norman ‘challenge [...] the notion that the laws of war serve to restrain or “humanize” war’, arguing instead that ‘through law, violence has been legimated’,4 while Mégret argues that treaties, statutes and sources of customary laws of war ‘enable, constitute and perpetuate’ war and have produced ‘the basic building blocks of the international grammar of violence’.5 Law, he proposes, is ‘just another name for fossilized theory that no longer sees itself as theory’. The theorising of the laws of war, Mégret argues, is necessary to ‘displace [...] the field’s confidence that it is “only law” and can shape our understanding of the laws of war as less immutable, progressive and benign than is typically accepted’.6 Elsewhere, he has also claimed that the laws of war are intimately tied up with the projection of masculine power, serving to provide a plausible, virulent masculinity that navigates ‘an impossible synthesis between violence and restraint’.7

Underlying these critiques is an intimation of the constructive power of language. These texts are not merely passive vehicles for a singular and authoritative voice of law, but are, rather, active players in the construction of principles and categories, at least some of which serve to legitimate rather than constrain geopolitical violence. They echo and bolster ideological forces outside the domain of law, not only available as a strategic asset by militaries8, but as a foundational ideological source for the maintenance of gendered power relations. Given the reliance of the laws of war on language, there is much to be gained by adding theoretical and methodological resources from linguistics to these critical accounts.

While linguistic engagement with legal discourse has a reasonably long history,9 the intersections are more recent in the field of the international laws of

---

6 Ibid 778.
war. The texts of the international laws of war are a natural site of collaboration between linguistics and law: they are made out of language, which gives a natural role for linguistics in their understanding, and they offer to linguistics the challenge of understanding and explaining their power to set the terms of legal geopolitical violence. No doubt part of the power of these texts of the international laws of war derives from the fact that they are interwoven with the threat of violence on which state sovereignty is based, but it would be misguided to think that what Althusser calls the ‘Repressive State Apparatus’ could function in a semiotic vacuum.

The organised violence that has proliferated in modernity has, as Malešević argues, been enabled by its ideological underpinnings. As such, these texts have not just a legal but an ideological role, shaping collective human thought and individual and institutional behaviour. While part of their own kind of text practice, the meanings of these texts are not confined to the legal field. To understand both their power and their contradictions, we need to understand the specific nature of these texts as a kind of text form, or ‘register’, and inquire into their semantic patterning via various linguistic approaches.

The paper is divided into three sections. The first examines the international laws of war as a ‘register’. The term ‘register’ denotes language variation based on context and recognises social and cultural context as the dominant shaping force of language both as a generalised semiotic system and in each specific use of language. As we discuss below, modern international laws of

---


14 A gaming satirical news site, Hard Drive, published an article in 2021 stating the the ‘New Geneva Conventions’ was now outlawing the killing of a player away from their keyboard. While these texts are developed by the international legal community, their meanings are open to recontextualisation.

war are based on a new kind of register, i.e., a new way of using language beginning, more or less, with the Paris Declaration of 1856.\textsuperscript{16} A new register is a new semiotic ecosystem, providing a semiotically constructed home in which to develop and maintain the meanings that give this textual practice its unique identity. These texts provide the conditions for the creation of the key principles of the international laws of war. Such principles are no more and no less than patterns of wording and meaning that such texts have created and which they sustain.

Section two of the paper treats the international laws of war texts as a ‘corpus’,\textsuperscript{17} a coherent body of searchable texts. It draws on corpus linguistic methods to inquire into their dominant semantic patterns. Using our expertise in linguistics, we constituted this corpus by incorporating all the texts included by the International Committee of the Red Cross in their International Humanitarian Law (IHL) treaties database,\textsuperscript{18} a set of 110 texts at the time of writing. It includes all current international treaties pertaining to the governance of war or ‘armed conflict’, treaties no longer in force, and documents, such as the Lieber Code, considered to have been formative in the development of this field. We demonstrate how key techniques in corpus linguistics, such as word frequency, concordance (viewing a particular word in the context of the words on either side of it), and collocations (a summary of the typical words that co-occur with a word under examination), can be used to explore significant lexical patterns in this body of text.

Finally, while corpus techniques enable certain reading practices over a large data set, texts are not simply bags of words: they are intricately structured, and subject to the grammatical systems of the language in which they are expressed. Section three of the paper provides a third linguistic angle on the international laws of war. It explores two extracts from the Rome Statute,\textsuperscript{19} the treaty that established the International Criminal Court, described by former UN Secretary-General Kofi Annan as ‘one of the finest moments in the history of the United Nations’ and ‘a giant step forward in the march towards universal human rights and the rule of law’.\textsuperscript{20} In this third section, we employ

\begin{itemize}
\item \textsuperscript{16} Paris Declaration Respecting Maritime Law (adopted 16 April 1856, entered into force 18 April 1858) 115 CTS 1.
\item \textsuperscript{17} A McEnery and A Hardie, \textit{Corpus Linguistics: Method, Theory, Practice} (Cambridge University Press 2012).
\item \textsuperscript{18} See <https://ihl-databases.icrc.org/ihl>.
\item \textsuperscript{19} Rome Statute of the International Criminal Court (adopted 17 July 1998, entered into force 1 July 2002) 2187 UNTS 90 (‘Rome Statute’).
\end{itemize}
a lexicogrammatical analysis, based on Halliday and Matthiessen’s work, to explore the meanings that grammatical patterns construe and how a short piece of grammatical structure can interweave deeply contradictory meanings. These three sections are intended to provide avenues for further interdisciplinary collaborations between law and linguistics.

2 The International Laws of War as ‘Register’

Open multilateral treaties, treaties to which any state can become a party, are a relatively new text form which appeared in the mid-19th century. It has been argued that the first open multilateral treaty was the 1856 Paris Declaration on Maritime Law, which was the first to attempt to regulate the behaviour of states at war. According to Roberts, it is not simply that the Paris Declaration is the first multilateral laws of war treaty: it is the first multilateral treaty.21 At a mere 336 words, the Paris Declaration introduces the idea of presenting, in written form, a set of four points, or ‘maxims’ – the only use of this word across the subsequent nearly 170 years of making of international laws of war – which are agreed upon by those states who declare themselves signatories.

Although this little text foreshadowed the emergence of what in linguistics is called a new ‘register’ – i.e., a functional variety of language – the idea of ‘laws of war’ well precedes it. Texts published from as early as 1535, available via the Early English Books Online (EEBO)22 corpus, show the expression ‘laws of war’ to be already part of public discourse, perhaps as a function of ordinances on the conduct of war promulgated by medieval kings, or tracts of medieval jurists23. Thus, the Paris Declaration enters a stage already prepared by an existing notion of the ‘laws of war’, but begins a process of textual and legal innovation, by creating a space for international negotiations on the behaviour of states ‘in time of war’, to quote the preamble. The Paris Declaration was the first step of many towards the creation of a body of international laws of war. Two produced soon afterwards – the 1863 Lieber Code24 and the 1864 Geneva

---

21 Ironically, as he argues, the ‘basic legal art form of the past 160 years or more’ actually emerged in the field of the laws of war, despite the fact that it is ‘undoubtedly a particularly difficult legal subject area’; AS Roberts, ‘Foundational Myths in the Laws of War: The 1863 Lieber Code and the 1864 Geneva Convention’ (2019) 20 Melbourne Journal of International Law 158, 164.
22 See <https://ota.bodleian.ox.ac.uk/repository/xmlui/handle/20.500.12024/A00381>.
24 General Orders No 100: Adjutant-General’s Office, Instructions for the Government of Armies of the United States, in the Field. Prepared by Francis Lieber, promulgated as General
Convention\textsuperscript{25} – would provide ‘foundational stories’ about the emergence and function of the international laws of war.\textsuperscript{26}

To understand the cultural and linguistic shift entailed in the emergence of the language of the ‘laws of war’, ‘register’ enables us to identify the features of the new text forms and appreciate those that may have been taken for granted. Registers have three key defining parameters, described as, (1) the nature of the social activity the text is engaged in (referred to as its ‘field’); (2) the nature of the relationships that bring the interactants together (referred to as ‘tenor’) and (3) the nature of the language contact (referred to as ‘mode’).\textsuperscript{27} Every text is some kind of register, configured by these three principles, because language itself evolved through these same principles of human interaction. These parameters are inescapable features of every act of meaning, with all uses of language being responsive to these three dimensions of context.

By using these linguistic concepts, some significant features of the international laws of war as a register can be identified. For example, the concept of tenor (the nature of the interactant relations) reveals that the international laws of war provided the conditions for the emergence of a new kind of discursive speaker role. These texts, starting with the Paris Declaration, acted as a vehicle for the ‘voice’ of an open collection of (emerging) nation state entities. Within two years of the Declaration’s promulgation, the text had attracted 41 signatories.\textsuperscript{28} Language was being used to perform a new kind of interaction, and in the process, the voice of a globalising, institutionalised speaker was constructed.

In the Paris Declaration, this speaker position was known as a collection of ‘plenipotentiaries’, a Latin term attested at least as early as the mid-17\textsuperscript{th} century in English.\textsuperscript{29} It appears in international war law until the 1986 Convention on the Rights of the Child,\textsuperscript{30} although its use past 1929 is very minimal.

\textsuperscript{25} Convention for the Amelioration of the Condition of the Wounded in Armies in the Field (adopted 22 August 1864, entered into force 22 June 1865) 129 cts 361.
\textsuperscript{26} Roberts (n 21).
\textsuperscript{27} Halliday and Hasan; Hasan (n 15).
\textsuperscript{28} Roberts (n 22) 164.
\textsuperscript{29} “plenipotentiary, adj. and n.” (\textit{OED Online}, March 2022) <www.oed.com/view/Entry/145683>.
1863 Lieber Code introduces the terms ‘Contracting Powers’ and ‘Contracting Parties’, the former falling out of use by 1934, and the latter last used in the 2006 International Convention for the Protection of All Persons from Enforced Disappearance. From 1968, the authority of a ‘Contracting Power’ would share the field with the term ‘States Parties’, which remains the current term to refer to document signatories.

This changing terminology – from the concept of an individual person invested with the authority to speak on behalf of a ruler, to an abstract, bureaucratic, contractual authority eventually based in statehood – reflects the long-term sociological process of bureaucratic rationalisation where ‘personal ties and nepotistic relationships become slowly but steadily replaced with impersonal rules and bureaucratic regulations’ – see Figure 1.

While Roberts suggests that the emergence of international cooperation should have been quite unlikely in such a contentious field as that of war, Malešević, argues that, in fact, it the origins of modern bureaucratisation, with its newfound force and reach, is with the military, quoting Weber’s point that ‘military discipline gives birth to all discipline’. The Paris Declaration can be seen as heralding a process by which an international bureaucratic authority emerges to set the ‘rules’ of war. It marks a milestone in a sociohistorical process Malešević calls ‘the bureaucratisation of coercion’, which he argues develops in parallel with war becoming ‘a total event involving millions of mobilised and ideologised citizens bent on the destruction of entire societies deemed to be enemies’.

The ‘voice’ of States Parties (and their predecessors) could only have been enabled and constructed via writing, a point which brings in the concept of ‘mode’ – the means by which language enables a speaker and addressee to come into contact. The texts of international war law are densely written texts created in the standard languages of the nation state. Their grammatical structures and patterns stand in stark contrast to the grammar of the spoken discourse of everyday life and echo the semantic style of other registers typical of modernity, such as the discourses of science and government administration. These registers – described as ‘inherently displaced’, that is, registers

---

31 Malešević (n 13) 5.
32 ibid 6.
33 Malešević (n 13) 7.
Figure 1. Terms describing interactants in texts of the international laws of war 1856–2017, including timeframe of usage and normalised frequency (number of tokens per million words, or wpm, discussed below) across the data.
where the reception of meaning is displaced in time and space from its production – have a grammatical complexity known as ‘grammatical metaphor’.36

Grammatical metaphor describes the grammatical tendencies in registers, such as in law, to favour nominal modes of representation: process being represented as a noun. One example is the concept ‘military necessity’, where the verb ‘to need’ has become a noun expressing an abstract thing. This results in reification and abstraction away from material, concrete, everyday experience. As Halliday argues, ‘grammatical metaphor allows users of standard languages to ‘hold’ […] the world still, making it noun-like (stable in time) while it is observed, experimented with, measured and reasoned about’.37 These linguistic patterns and habits allow communities to create ‘a parallel semiotic universe of virtual things, especially virtual entities and virtual processes’.38

The ‘virtual phenomena’ created through grammatical metaphor exist only on a semiotic plane: they have no direct material analogue, and are entirely dependent on their definitions. For example, the key principles of the international laws of war, summarised in Crawford and Pert, are ‘distinction’, ‘military necessity’, ‘proportionality’, ‘the prohibition on causing unnecessary suffering and superfluous injury’, ‘neutrality’, and ‘humanity’,39 all instances of grammatical metaphor, all relying on the grammatical shift into abstract nouns, a grammatical shift from meanings more directly or concretely expressed through other grammatical structures.

This feature can be seen as soon as we attempt to ‘translate’ these concepts into everyday language: the principle of ‘distinction’ can be translated to ‘you have to distinguish between ordinary people and the people who belong to the army of a country’; ‘the prohibition on causing unnecessary suffering and superfluous injury’ could be ‘you are not allowed to make people suffer too much and you cannot injure them too much either’. Translated, these concepts lack the mystique and prestige of the abstract form, the capacity to have the status of a legal principle. These concepts depend on the grammar of densely


written text for their existence and identity. This is the meaning of ‘semiotic constructs with no material analogue’, and these constructs came into being through the emergence of the international laws of war register.

In summary, the texts of the international laws of war herald a new register, in which a newly forming speaker position (i.e., tenor) – an abstract, international, bureaucratic voice, with the force of nation state militaries – projects itself through the densely crafted and highly elaborated written form, dependent on grammatical metaphor (i.e., mode). The function of this discourse (i.e., field, or nature of the social activity) as these texts accumulate over time, is, as we have already seen, subject to debate and interpretation. While these texts have an overtly stated purpose, often articulated in a preamble, the complex combination of tenor and mode ensures that these texts can be home to discordant meanings and internal contradictions. More detailed analysis of the language of the texts of the international laws of war is one means to bring out such contradictions.

3 The International Laws of War as Corpus

Despite the contest over precisely what semiotic work they perform, the texts of the international laws of war are functionally similar enough to be treated as a ‘corpus’, that is, as a searchable body of linguistic data. For our purposes, we have assembled all the texts hosted by the International Committee of the Red Cross in its IHL database.40 While the International Committee of the Red Cross website, which hosts these documents, has improved its search functionality,41 these texts are now available in a search interface with the full functionality of linguistic corpus technology via the Sydney Corpus Lab42 under the title of ‘Macquarie Laws of War Corpus’ (MQLWC), reflecting the institution (Macquarie University) where this corpus development work took place.

40 See <https://ihl-databases.icrc.org/ihl>.
41 See https://www.icrc.org/en/document/searching-ihl-databases-icrc. The site covers treaties, customary law, and documents associated with national implementation, and includes documents in seven languages. The site currently only allows a user to determine if a word or phrase appears in the available documents.
42 The Sydney Corpus Lab is a corpus website directed by Professor Monika Bednerak at the University of Sydney. By requesting a free account, any user can access available corpora at this site: https://sydneycorpuslab.com. For details on the MQLWC, see Lukin and Araujo e Castro (n 10).
The corpus is based on the 110 texts (English only; the corpus excludes Annex II to Protocol Additional to the Geneva Conventions 1949, as it consists of a template only) hosted in the International Committee of the Red Cross IHL database, a corpus totaling c. 392K words, beginning with the Paris Declaration in 1856 and extending up to 2019 when a further amendment to the Rome Statute was made. In this format, this body of international law can now be easily searched with basic corpus linguistic techniques such as word frequencies, concordances (where a word of interest is viewed in relation to words both to the left and right of this word in a corpus) and collocations (i.e., the typical words that co-occur with a specific word), providing the opportunity to study these texts in new ways. In addition, the whole corpus can be downloaded from the Sydney Corpus Lab for offline processing via other popular corpus software tools. This new access to these texts opens them up for a range of inquiries. Here we illustrate this potential with a brief discussion of the data using some corpus linguistic techniques.

Starting with word frequencies, Table 1 provides an example of the kind of information that can be extracted from this body of law. It lists, in order of frequency, the top 20 'open system' lexical items across the MQLWC; that is, the list of top frequency words when grammatical words (articles, such as 'the'; prepositions, such as 'in' or 'of'; and conjunctions, such as 'and', among others) are excluded. The table specifies both 'raw' and 'normalised' frequencies, the first being actual number of occurrences in the data, and the second being a measure of frequency per quantum or words, using a standard measure of 'words per million', or 'wpm'. Normalised frequency allows us to compare patterns in one data set with another. In addition, the metric of text dispersion shows how many texts in the corpus include these words.

This list includes some high frequency items that reflect the bureaucratic process itself of 'states' 'parties' signing up to 'international' 'conventions', which are constituted by 'articles' (and where 'art' is an abbreviation for this term), reflecting that these texts are part Malešević’s 'bureaucratization of

43 Different corpus systems have different ways of defining words or 'tokens'. This count, based on the cqweb system, includes punctuation items.
45 Analysing word frequencies provides a measure of dominant word forms and, to some degree, dominant concepts in a corpus.


<table>
<thead>
<tr>
<th>Rank</th>
<th>Lexical item</th>
<th>Raw frequency</th>
<th>Normalised frequency (wpm)</th>
<th>Text dispersion (n/total n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>article</td>
<td>2722</td>
<td>7667</td>
<td>86/100 (78%)</td>
</tr>
<tr>
<td>2</td>
<td>convention</td>
<td>1719</td>
<td>4842</td>
<td>77/110 (70%)</td>
</tr>
<tr>
<td>3</td>
<td>parties</td>
<td>1645</td>
<td>4638</td>
<td>85/110 (76%)</td>
</tr>
<tr>
<td>4</td>
<td>state</td>
<td>1544</td>
<td>4349</td>
<td>68/110 (68%)</td>
</tr>
<tr>
<td>5</td>
<td>art</td>
<td>1407</td>
<td>3963</td>
<td>80/110 (73%)</td>
</tr>
<tr>
<td>6</td>
<td>war</td>
<td>1401</td>
<td>3946</td>
<td>79/110 (72%)</td>
</tr>
<tr>
<td>7</td>
<td>party</td>
<td>1344</td>
<td>3785</td>
<td>63/110 (57%)</td>
</tr>
<tr>
<td>8</td>
<td>international</td>
<td>1154</td>
<td>3250</td>
<td>90/110 (81%)</td>
</tr>
<tr>
<td>9</td>
<td>states</td>
<td>1047</td>
<td>2949</td>
<td>65/100 (59%)</td>
</tr>
<tr>
<td>10</td>
<td>power</td>
<td>970</td>
<td>2732</td>
<td>52/110 (47%)</td>
</tr>
<tr>
<td>11</td>
<td>present</td>
<td>846</td>
<td>2383</td>
<td>78/110 (71%)</td>
</tr>
<tr>
<td>12</td>
<td>prisoners</td>
<td>832</td>
<td>2344</td>
<td>34/110 (31%)</td>
</tr>
<tr>
<td>13</td>
<td>military</td>
<td>827</td>
<td>2329</td>
<td>66/110 (60%)</td>
</tr>
<tr>
<td>14</td>
<td>conflict</td>
<td>791</td>
<td>2228</td>
<td>42/110 (38%)</td>
</tr>
<tr>
<td>15</td>
<td>court</td>
<td>752</td>
<td>2118</td>
<td>37/110 (34%)</td>
</tr>
<tr>
<td>16</td>
<td>persons</td>
<td>712</td>
<td>2005</td>
<td>67/110 (61%)</td>
</tr>
<tr>
<td>17</td>
<td>paragraph</td>
<td>675</td>
<td>1901</td>
<td>68/110 (62%)</td>
</tr>
<tr>
<td>18</td>
<td>provisions</td>
<td>675</td>
<td>1901</td>
<td>78/110 (71%)</td>
</tr>
<tr>
<td>19</td>
<td>powers</td>
<td>665</td>
<td>1873</td>
<td>57/110 (52%)</td>
</tr>
<tr>
<td>20</td>
<td>contracting</td>
<td>664</td>
<td>1870</td>
<td>60/110 (54%)</td>
</tr>
</tbody>
</table>

coercion.\textsuperscript{47} Setting these words aside for now, then ‘war’, ‘power’, and ‘military’ can been seen as very dominant concepts across this body of text, empirical findings which can provide evidence for debates about the preoccupations within this register.

\textsuperscript{46} These frequencies were derived using Lancsbox: V Brezina, P Weill-Tessier, and A McEnery, \#LancsBox v. 5.x. [software] (2020). Available at: <http://corpora.lancs.ac.uk/lancsbox>.

\textsuperscript{47} Malešević (n 13).
While each of these terms merit further investigation, for our purposes we consider in more detail the use of the word ‘military’ in this corpus. The mirror image of ‘military’, the word ‘civilian’, does not appear in the top 20 most frequent terms in this data. While ‘military’ ranks 13th, ‘civilian’ ranks at 61st, and appears at less than half the frequency of ‘military’ (359 for ‘civilian’; 827 for ‘military’), and is found in approximately 50% fewer texts than ‘military’ (34% text dispersion for ‘civilian’; 60% for ‘military’). A further measure of its frequency can be sought by comparing the frequency of the word ‘military’ in this corpus to a variety of other corpora to consider whether ‘military’ is particularly frequent in the international laws of war. This is likely to be the case, but a comparison with other data not only provides the evidence for this hypothesis, but can also provide a calibration of the greater frequency of ‘military’ in the laws of war.

To do this, we selected relevant comparative corpora, such as the British National Corpus (bnc: 100 million word cross-register corpus of British English from early 1990s); AmE06 (1 million word cross-register corpus from American English between 2005–2007); BE06 (1 million word cross-register corpus from British English between 2005–2007); News on the Web corpus (>15 billion words of news data in English across 20 countries from 2010 to current).48 Figure 2 plots the normalised frequency for ‘military’ across these five corpora, with the MQLWC clearly showing a strong bias towards the word. The average normalised frequency across the comparator corpora for the word ‘military’ is 157 wpm. With the normalised frequency of ‘military’ in the laws of war corpus being 2329 wpm, the comparison shows ‘military’ to be nearly 15 times more frequent in the laws of war.

Again, these findings are not surprising: but if claims are being made that this discourse form is not restraining but instead enabling violence, then the high frequency of ‘military’ may be indicative of a legitimising labour that these texts are engaged in. A further step methodologically is to consider the ‘collocates’ of a lexical item – that is, the words which typically go with (on either side of) a word of interest, for instance, taking five words either side of the lexical item. Collocates can be determined simply on the basis of frequency, but other measures allow collocates to be determined by a combination of frequency and mutual expectancy (the probability that two lexical items will

typically co-occur). Table 2 shows the top 20 collocates for the word ‘military’ in the corpus of the international laws of war (leaving out function words).49

These collocates give ‘military’ a rational underpinning, by associating it with lexical items that imply organisation and planning, such as ‘objective/s’, ‘operations’, ‘objects’ and ‘purposes’. Military power is authorised by its association with the words ‘authority/ies’. The conjunction of ‘military’ and ‘action’ creates a large category of behaviour without any specificity, thereby standing back from the embodied violence of the specific ‘actions’ that militaries engage in, which include invading, bombing, killing, injuring, destroying, etc. ‘Action’ is instead a neutral, catch-all phrase that avoids the negativity of other likely word selections.

The conjunction of ‘military’ with ‘necessity’ can also be observed in this list, in addition to various synonyms such as: ‘advantage’, ‘imperative’ and ‘considerations’. As Lukin and Araújo e Castro note, the international laws of war are the semiotic ecosystem in which the concept of ‘military necessity’ was

---

49 See V Brezina, T McEnery, and S Wattam, ‘Collocations in Context: A New Perspective on Collocation Networks’ (2015) 29 International Journal of Corpus Linguistics 139. The collocates in Table 3 are based on the following collocation parameters: m13 (9); L5-R5;C5; NC5, that is, a minimum threshold of m13 of 9; a collocate window of 5 left and 5 right; a minimum collocate frequency of 5 and a minimum frequency of collocation of 5.
created, and has been maintained.\textsuperscript{50} It is inside of this body of law that ‘military necessity’ is, albeit very briefly, and in a fundamentally circular fashion, defined for legal purposes. The concept almost does not exist outside of this paradigm context (i.e., it appears as a marginal concept in data sets such as the BNC, the Google Books Corpus, and the News on the Web corpus). The international laws of war provide the ‘parallel universe’ in which ‘military necessity’ can operate and have plausibility.

\begin{table}
\centering
\caption{Top 20 collocates of ‘military’ in MQLWC}
\begin{tabular}{llll}

\hline

Lexical item & M13 & Freq (collocation) & Freq (corpus) \\
\hline
1 & objectives & 21.40 & 89 & 109 \\
2 & operations & 20.21 & 83 & 202 \\
3 & objective & 18.88 & 40 & 57 \\
4 & necessity & 18.80 & 46 & 91 \\
5 & aircraft & 18.49 & 72 & 433 \\
6 & advantage & 18.25 & 32 & 45 \\
7 & authorities & 17.62 & 48 & 235 \\
8 & military & 17.05 & 64 & 827 \\
9 & action & 16.78 & 30 & 103 \\
10 & imperative & 16.25 & 17 & 27 \\
11 & direct & 16.16 & 26 & 103 \\
12 & authority & 16.15 & 31 & 176 \\
13 & permit & 16.07 & 24 & 86 \\
14 & objects & 15.94 & 27 & 134 \\
15 & art & 15.86 & 58 & 1407 \\
16 & purposes & 15.85 & 31 & 216 \\
17 & anticipated & 15.84 & 13 & 16 \\
18 & considerations & 15.70 & 14 & 22 \\
19 & concrete & 15.54 & 11 & 12 \\
20 & establishments & 15.53 & 19 & 62 \\
\hline
\end{tabular}
\end{table}

\textsuperscript{50} Lukin and Araujo e Castro (n 10).
Linguistically, ‘military necessity’ is only the conjunction of a classifier/adjective (‘military’) and a noun (‘necessity’). While all words are linguistic entities, ‘military necessity’, like anything we refer to as a ‘principle’, exists only within a discursive framework. This is not the case for all entities inside the documents of the international laws of war. One strand in these texts has been around attempts to limit war’s lethal technology, such as exploding bullets under 400 grams, landmines, and chemical and nuclear weapons. All of these have a very real and material existence outside the discourse of international law. This does not mean that language is simply a mirror to them – language is always exercising its organising and categorising power. But, ‘military necessity’ is, as noted above, a ‘virtual entity’: it has no existence outside of language because it relies, as all such principles do, on grammatical metaphor.

Let’s recall its original, and notably circular, definition in Article 14 of the Lieber Code where ‘military necessity’ is defined as consisting ‘in the necessity of those measures which are indispensable for securing the ends of the war, and which are lawful according to the modern law and usages of war’. The definition has no special technical terms: ‘necessity’ is simply a nominalised form from the verb ‘to need’, with the conjunction of the term ‘military’ acting to grant the illusion of a specialised sub-category of ‘necessity’. If we attempt to ‘de-metaphorise’ the concept, that is, to translate the nominalised form into the grammar of our everyday interactions, we end up with an explanation in these terms: ‘militaries need to do things’. In these terms, the concept loses the prestige bestowed on it through its nominal form. It becomes vague and starts to resemble an ambit claim. But the concept is validated by its constant repetition in these texts, and by the many validating collocates that come with the term ‘military’.

In this section, we have sought to exemplify the potential of corpus linguistics techniques, with a small investigation of the dominant words in this corpus, followed by some more detail on the word ‘military’. While there is much to be elaborated on here, we turn now to complement the two approaches previously explored with a final angle from linguistics, namely the study of lexicogrammatical patterns in this data. Given the intricacy of grammar, we take one small section from the 1998 Rome Statute, to consider the meanings made through lexicogrammatical selections and patterning.

4 Grammatical Patterning in the Texts of the International Laws of War

The texts of the international laws of war, like all others, are subject to the semiotic forces which are involved in creating a text, in which grammatical
systems are the foundation. For any act of meaning to perform its necessary work, it must have three interwoven strands of meaning, which derive from what Halliday calls the ‘metafunctional’ organisation of language. All languages, he argues, have three modes of meaning which texts organise through grammatical structures. These three strands of meanings are: experiential meanings, which enable speakers to construe our experience of the world; interpersonal meanings, through which we enact our interpersonal relationships in all their complex variability; and textual meanings, the mode of meaning which brings experiential and interpersonal meanings into a coherent flow of discourse. These modes of meaning reflect the three principles of interaction (field, tenor and mode) outlined earlier.

All texts, legal or otherwise, traverse the grammatical systems of a language. In the creation of a text, speakers and writers make selections from these grammatical systems, and in the process a text’s potential meanings are configured. For the purposes of illustrating the ideological potential of grammatical patterns in the laws of war, some examples from a small but significant extract of the Rome Statute are examined below.

As previously mentioned, this treaty established the International Criminal Court and was heralded by Kofi Annan as a very significant step in the history of the UN. The Rome Statute proscribes and criminalises the actions labelled ‘war crimes’, ‘crimes of aggression’, ‘crimes against humanity’ and ‘genocide’. As much as it might seem that such crimes are self-evident, within a legal context they require definition to indicate their scope. Only language can provide such definitions: what makes an action a ‘war crime’ is the meaning that can be attributed, via language, to some forms of action and not others. The act of defining ‘war crimes’ involves construing through wording the behaviours that will be attached to this crime, and such a definition cannot be achieved without choosing grammatical options from these grammatical systems. Language is not reflecting, but constructing these categories.

Tables 3 and 4 provide two example clauses from Article 8 of the Rome Statute, which carries the burden of defining the meaning of the legal concept ‘war crimes’. These examples are analysed with respect to two grammatical systems of Halliday’s ‘experiential’ function of language: the systems of thing type and process type. All English language text construction is

52 Annan (n 20).
53 The grammatical analysis in this paper is based on MAK Halliday and CMIM Matthiessen, An Introduction to Halliday’s Functional Grammar: Fourth Edition (Arnold 2014). The small caps formatting is a standard for the name of a grammatical system.
**Table 3**  Example 1 – Article 8 (b) (xx)\(^{54}\)

<table>
<thead>
<tr>
<th>Verbal group</th>
<th>Nominal group</th>
</tr>
</thead>
</table>
| Employing    | weapons, projectiles and material and methods of warfare \([\text{which are of a nature \([\text{to cause superfluous injury or unnecessary suffering}]\) or which are inherently indiscriminate in violation of the international law of armed conflict, \(\|\) provided that such weapons, projectiles and material and methods of warfare are the subject of a comprehensive prohibition \(\|\) and are included in an annex to this Statute, by an amendment in accordance with the relevant provisions set forth in articles 121 and 123;}\]|}

**Table 4**  Example 2 – Article 8 (b)(iv)\(^{56}\)

<table>
<thead>
<tr>
<th>Verbal group</th>
<th>Nominal group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intentionally launching an attack in the knowledge ([\text{that such attack will cause incidental loss of life or injury to civilians or damage to civilian objects or widespread, long-term and severe damage to the natural environment ([\text{which would be clearly excessive in relation to the concrete and direct overall military advantage ([\text{anticipated}])]}]})</td>
<td></td>
</tr>
</tbody>
</table>

---

54 Rome Statute (n 19).

55 Double square brackets ‘\(\ldots\)’ denote an embedded clause, i.e. a clause which functions as a constituent element, or part of a constituent element, within a clause; two vertical lines ‘\(\|\)’ denote a clause boundary. See MAK Halliday and CMIM Matthiessen, *An Introduction to Halliday’s Functional Grammar: Fourth Edition* (Arnold 2014).

56 Rome Statute (n 19).
subject to the options of these grammatical systems: all texts have some kinds of entities or ‘things’ that construct content meanings in a text (THING TYPE) and all texts include lexical verbs or processes (PROCESS TYPE). The analysis of patterns in these grammatical systems will enable us to see a little more deeply into the semiotic mechanics of the ways in which the violence of war is constructed in this body of law.

Example 1 in Table 3 consists of two constituent parts: a ‘verbal group’ construing the process, and a ‘nominal group’ construing an entity or ‘thing’ pertinent to this context. Nominal groups are responsible for creating what humans want, need, or dream up. A ‘thing’, grammatically, can be short and sweet, such as ‘poison or poisoned weapons’, a simple and relatively unelaborated flow of words. Presumably the lack of elaboration in the nominal group ‘poison or poisoned weapons’ is because the harm of such weapons appears to

<table>
<thead>
<tr>
<th>Grouping</th>
<th>Verbal Group</th>
<th>Nominal Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Function</td>
<td>Process: Material</td>
<td>Goal</td>
</tr>
</tbody>
</table>

---

57 Rome Statute (n 19)
58 While the terminology of ‘verb’ and ‘noun’ is likely to be more familiar, these concepts are not robust enough to capture the semiotic work done by the structures Halliday calls verbal and nominal groups, and which have a variety of functional roles in a clause.
59 Rome Statute (n 19) art 8(b)xvii.
be uncontested, so no further elaboration, conditions or caveats are required to understand this particular category of thing.

Legal texts are home to both elaborated and abstract things, and the examples in Tables 3 and 4 each show clauses in which one constituent element (the nominal group) is highly elaborated. The length of these structures is one key indicator of this complexity. The presence of long, complex nominal groups in a legal text is not in any way a surprise: but neither are such structures arbitrary or to be taken-for-granted. These ‘things’ in this text are expressed through nominal groups with extensive elaboration or ‘post-qualification’ (i.e., wording that follows the noun, and which is required to clearly specify the category of thing being invoked in the text).

Example 1 in Table 3 consists of two parts, a simple verbal group (only one word is required for this constituent to do its job), and a long and complex nominal group which construes a highly specific entity in the text. With so much intricacy in a nominal group structure like this, it is perhaps also unsurprising that we find within this structure contradictory semantic currents,

---

60 Rome Statute (n 19)
described by Mégret as the ‘impossible synthesis between restraint and violence’. Example 1 starts by heading towards the restraining of some particular war-like behaviour, but before we get too far into this clause, we encounter the placing of limits on these limits. The grammar of this clause gives with one hand, and then takes with the other.

Table 7 shows the internal group structure analysis of the ‘thing’ or nominal group element of the clause. The lack of determiner (e.g., the, some, those, these etc.) or any other premodification means this instance includes all members of the category ‘weapons, projectiles and material and methods of warfare’,

---

Table 7  Nominal group structure of Goal element in Examples 1 and 2

<table>
<thead>
<tr>
<th>Example 1</th>
<th>null</th>
<th>weapons, projectiles and material and methods of warfare [of warfare]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>[[which are of a nature [[to cause superfluous injury or unnecessary suffering]]</td>
</tr>
<tr>
<td>Example 2</td>
<td>The Knowledge</td>
<td>[[that such attack will cause incidental loss of life or injury to civilians or damage to civilian objects or widespread, long-term and severe damage to the natural environment [[which would be clearly excessive in relation to the concrete and direct overall military advantage [[anticipated]]]];</td>
</tr>
</tbody>
</table>

Determiner     Thing     Post-qualifier

---

creating the appearance of a broad all-encompassing sweep of prohibition. Eight words into this structure, the post-qualifying element (the words following the central noun) begins, adding a further 72 words, and in the process considerably reduces the scope of reference of this nominal group. This post-qualifier begins by reducing the category ‘weapons, projectiles and materials and methods of warfare’ to those which cause ‘superfluous’ or ‘unnecessary’ injuries and suffering.

These adjectives are relative: they can only be analysed from a particular, and potentially variable, benchmark which is unspecified. This clause is joined to the next by the conjunction ‘or’, providing an alternative means by which this category of ‘weapons … etc’ is curtailed, i.e., if 1. they are shown to be ‘indiscriminate’ in a way that violates international law (again, a contestable principle) and 2. if they have been subject to comprehensive prohibition, and 3. if they are listed in the annex to the Statute. In summary, the only weapons, projectiles and methods of warfare prohibited are those ‘which are \(w\), or which are \(x\), provided that they are \(y\) AND that they are \(z\). This is quite a hurdle to have to pass for a weapon or method of war to be prohibited. To date, no form of weapons, projectiles or materials and methods of warfare has been so listed in an Annex to this Statute.62

Example 2 (Table 4) is a variation on this theme, although now the complex thing being constructed is the state of mind or ‘knowledge’ of the belligerent. This clause in Article 8 criminalises only those attacks carried out ‘intentionally’, and ‘in the knowledge that … etc’ – a point echoed in Article 30 of the Statute (‘Mental Element’) which states the requirement of knowledge and intent for criminality – essentially, the long-held common law principle of \(\text{mens rea}\).63 As such, two of the four constituents of this clause are devoted to giving weight not to an action or its consequences, but to the consciousness brought by the belligerent to this action. The second of these constituents, as shown in Table 5, is a lengthy discourse spelling out what kind of ‘knowledge’ results in an attack being criminalised. It is not the fact of, or scale of, the incidental killings, woundings or suffering caused. Instead, such loss, injury or damage must be shown to be ‘excessive’ relative to the ‘military advantage [[ anticipated]]’.

---


63 According to Ausserladscheider Jonas (2021), the Rome Statute is the only international instrument with a definition of \(\text{mens rea}\) (Article 30), although it is also present in customary international law. See L Ausserladscheider Jonas, \(\text{Individual Criminal Responsibility for Financing Entities Involved in Core Crimes}\) (Brill 2021) 40, 40–44.
As we have seen before, a lexical item like ‘excessive’ is relative to some particular, and potentially variable, threshold – it is not a self-evident quantifier. Furthermore, the thing to which the threshold is relative is a highly abstract concept, namely ‘military advantage’, also known as ‘military necessity’ – as discussed above, a legal fiction which exists only within the confines of the international laws of war. Around this concept, we again can see two contradictory currents: the advantage has to be ‘concrete’ and ‘direct’, an attempt to try to ground this highly abstract concept in evidence. Yet, at the same time, the measure is not actual ‘military advantage’, but only that which was ‘anticipated’. In other words, the advantage has to be concrete and direct at least in the mind of the author of the attack prior to it taking place. This second example shows us, again, the grammar restraining violence on the one hand and enabling it with the other – all within the confines of one single element in one single clause.

As all the material and methods of warfare are indisputably lethal, Article 8 is attempting to provide some magical space in which this lethal technology can be used by some legal justification. For those who claim that these texts work to constrain violence, the text of the Rome Statute clearly shows itself attempting to regulate actions and place limits of weapons. In addition, the florid elaboration of entities in the text creates the appearance of attention to the detail of what is in and what is outside of this definition, as if international legal minds have striven for, and achieved, a profound degree of specificity. The elaboration seen in these two entities (but of which there are many more in the text) is reiterated at another scale: the list of war crimes in Article 8 is spread over 4 subsections with eight listed in 8(a), 26 in 8(b), 4 in 8(c), and 15 in 8(e), creating the impression of significant semiotic energy being devoted to precisely pinning down the scope of the category ‘war crimes’.

Language has indeterminacy as one of its key features, a function of the fact that it is a ‘theorizing of the human condition’ in which it manages ‘a massive reconciliation of conflicting principles of order – this being the only way of ‘semioticizing’ our complex ecosocial environment in a way that is favourable to our survival’. Some degree of indeterminacy is therefore unavoidable: but language allows its speakers and writers key choices that bring semantic consequences.

For example, the long list of proscribed actions under the umbrella of ‘war crimes’ have a key characteristic, which is also a consequence of a grammatical choice. An action verb can be transitive or intransitive. A transitive verb or

process construes an action that extends to another entity and impacts upon it. ‘Killing’ is transitive; ‘injuring’ and ‘wounding’ are transitive, ‘torturing’ is transitive; ‘attacking’, ‘destroying’, ‘pillaging’, ‘raping’, ‘displacing’, ‘conscripting’ and ‘starving’ are all transitive verbs. Again, this comes as no surprise – surely the very essence of the methods of war are actions which are construed as highly impactful. It is a key feature of the overall tone of Article 8 that it contains such a long list of transitive processes, whose impact is unequivocally negative.

Violence which is stigmatised requires that it be construed in specific and transitive terms, so that its consequences are made explicit. But the choice of transitive grammar and semantics is not simply reflecting these actions, it also constructs them since language always offers alternative words and wording. An example of such contrasting wording is found in the Article (2)(e), pertaining to torture. While torture is criminalised, it is only torture when defined as ‘intentional infliction of severe pain or suffering’. By contrast, if the same pain or suffering simply ‘arises’ as a consequence of ‘lawful sanctions’, this does not fit the definition of torture, even if such pain and suffering inheres in (and is therefore caused by) the belligerent action. Here, the line between legal and unlawful is resolved by a grammatical selection, the contrasting of the transitive verb ‘to inflict’ with the intransitive verb ‘to arise’. In the first, pain and suffering is ‘inflicted’, while in the second it merely ‘arises’, and in this construction, agency and causation is made invisible.

The use of transitive grammar and semantics generates a set of proscribable and illegitimate actions – and so this language is working to constrain forms of violence. The extensive list of such actions adds weight to this sense of constraint. Mégret has argued that the laws of war enact their semantics of restraint in a theatrical fashion, perhaps as a means to draw attention from the many ways this language is also legitimising violence. Since this theatricality has to be an artefact of language use in these texts, arguably it is produced through the seemingly exhaustive list of crimes, their articulation as transitive processes, the repetition and redundancy across the various sub-sections of Article 8, and the florid qualification of things.

One further feature of this text could be added here: the apparent intent to ensure all alternatives of some category are voiced. One of the textual measures of this final point is the pervasive use of the conjunction ‘or’. The frequency

---

66 Mégret (n 61).
of the conjunction ‘or’ in this text can be calibrated against the corpus as a whole, and this corpus against other corpora: Table 8 sets out the frequency of ‘or’ compared with the most frequent of all conjunctions, ‘and’ in the Rome Statute, compared with the BNC, the BE06 and the AmE06 corpora. In the comparator corpora, the ratio of ‘and’ to ‘or’ is more than 7 to 1. In the MQLWC, the figure is around 1.5 to 1, and in the Rome Statute, the frequency of ‘or’ is actually higher than ‘and’.

In this short analysis of a very small extract from the Rome Statute, we begin to see how effortlessly grammatical systems are recruited to both constrain and enable the violence of states. Across a variety of micro-grammatical choices exemplified here, lethal action is proscribed, but perhaps more importantly, is seen to be proscribed, while simultaneously, and within one and the same grammatical structure, such proscriptions are limited, opening the semantic spaces for forms of violence to be given the imprimatur of legality.

5 Conclusion

The aim of this paper has been to show how linguistic theory and method can assist in examining the kind of semiotic work that the international laws of war perform. Each of the three sections of the paper approached this aim in different ways: section 2 introduced the concept of ‘register’, a necessary linguistic concept to understand the way in which these texts are unique products of deep cultural forces, creating not only the voice of a bureaucratic international community, but a semiotic ecosystem for the development of the defining principles of the international laws of war.

In section 3, the potential of corpus linguistics to study this body of law was examined and illustrated with reference to the word ‘military’, a very dominant
concept in this discourse. Its very high relative frequency shows the international laws of war to be a semiotic environment which gives this concept a very big platform to accrue associations. An examination of collocations of this word showed these associations to be deeply legitimating. Finally, section 4 brought the focus to the intricate work of grammatical structure, and the ways that its patterns covertly generate semantic currents with significant ideological consequences.

These three linguistic angles on these texts show the role language has played in constructing this discourse form, furnishing the environment for the development of legal principles which define this field, and providing the lexicogrammatical patterns that both proscribe and legitimate the violence of war. Each of these demonstrations provides a starting point for further interdisciplinary collaborations to enhance understanding about the power of language in a context where its processes and products have such profound human consequences.