Chapter 4

War and Violence: Etymology, Definitions, Frequencies, Collocations

There is a time for everything, and a season for every activity under the heavens ... a time to kill and a time to heal, a time to tear down and a time to build ... a time to love and a time to hate, a time for war and a time for peace.

Ecclesiastes 3

4.1 Introduction

In Chap. 1, I made a number of predictions about the nature of the word war, based on the account by Malešević about the relationships between the sociological function of war and the legitimacy afforded war by its ideologies. These predictions include: that war will be a highly frequent and highly dispersed lexical item; that it will attract many collocations so that the category itself is effectively neutral; that both its denotative and connotative meanings will generate largely neutral or positive meanings and associations; and that the category war will be strongly insulated from lexical options that undermine these neutral or positive meanings. The focus of this chapter is to make visible some of the legitimating semantics of war, taking the lexical item itself as a point of departure. To make, let alone to test, these predictions involves first of all understanding the complexity of words. Words, or signs, capture differences. They are placeholders, points of intersection of a large number of classificatory dimensions (Halliday 2002). Their meanings depend on the signs that go around them, their contexts of use, their social and cultural environments. As Hasan argues, “there is a continuity from the living of life on the one hand right down to the morpheme on the other” (Cloran et al. 1996, 1). This linguistic fact ensures that word meaning is complex. Indeed, Saussure argued that setting out a complete inventory of the meanings contained in a word was inconceivable.

1Fontaine (2017) has drawn the happy analogy between the word and the TARDIS of Dr. Who fame, not only because it travels in space and time, but also because it is “bigger on the inside”.

2Saussure acknowledged the possible exception to his claim of very rare words, giving examples such as aluminium and eucalyptus (de Saussure 2006, 52).
Comparing a word such as iron to steel, lead, gold or metal, he contended, would be a “never-ending task”. To examine the meaning of mind, as opposed to soul or thought, “would quite simply be a lifetime’s work”. Saussure even suggests that it is “pointless” to seek “the totality of a word’s meaning” (de Saussure 2006, 52). No less is this the case for the lexical item war, a word that, when examined with appropriate theoretical insights and empirical resources, turns out to be “sociologically symptomatic” (Firth 1957b, 13). To test my claim about the insulation of war from semantically proximal but negative lexis, this chapter also analyses war’s alter ego, the lexical item violence.

Key concepts relevant to understanding how words mean have been set out in the linguistics of scholars such as Saussure, Vološinov, and Firth, among others (discussed in Chap. 2). In this chapter, I will begin the exploration of war and violence by reference to standard lexicographical resources, drawing on the Oxford English Dictionary (hereafter OED) as a starting point. I will also consider their etymologies, and their contrasting locations within the conceptual architecture of both the OED thesaurus and Roget’s thesaurus. In doing so, it will be necessary to see war and violence not simply as lexical items, but as “lexicogrammatical” ones. Words cannot be studied in isolation: their very existence is tied up with the syntagmatic nature of language, and by extension, the social contexts in which they are relevant. Words are meaningful because they form structures; and linguists since at least Saussure’s time have recognised that the separation of lexis from morphology or syntax is untenable. Grammatical meanings are integral to the processes by which lexical items function. In even a preliminary investigation of war, for instance, we confront the empirical fact that it is now only used as a nominal element: although previously able to be used as a verb to construe an unfolding process, that form of war is now obsolete. Instead, war is now only a nominal element, subject to the kinds of grammatical environments and relations afforded this particular grammatical category. Such a development cannot be arbitrary, but must be related to changing cultural practices.

From the examination of these terms in the dictionary and thesaurus, the next step is to consider the behaviour of these words in large corpora. With many large corpora now available, such as the British National Corpus, the Google Books Corpus, the corpus of the British Hansard, and the Corpus of Contemporary American English, it is possible to examine patterns of usage of such key terms across many genres, and across time spans. A central corpus for the discussion in this chapter will be the British National Corpus (or BNC). The corpus is multi-genereric, composed of a 90 million word balanced written corpus (i.e. with equal samples of a range of written registers) and a c. ten million word corpus of orthographically transcribed spoken discourse. The written registers include academic prose, fiction, newspaper discourse, unpublished letters, school and university essays. The spoken discourse section includes “context-governed” speech, i.e.

---

3 This term is problematic, as all use of language is a realisation of its context of situation (see Chap. 3). Here the term is used to differentiate discourse via a spoken channel other than naturally occurring conversation.
samples of dialogic speech from four general domains, labelled “educational and informative”, “business”, “public or institutional” and “leisure”. It also includes “demographically sampled” speech, i.e. speech from natural conversation by speakers sampled from different social classes and geographic locations, and with due regard for differences in gender and age. The BNC is the most representative corpus of a modern language currently available, though it is British English, and is synchronic (that is, sampled from roughly the same time period, the late twentieth century). Other corpora will be referred to from time to time, and will be introduced as necessary.

4.2 War in the Dictionary and Thesaurus

The Oxford English Dictionary traces the history of the modern word *war* to a late Old English word (*wyrrlewerre*, c. 1050), coming into Old English from North Eastern Old French (*werre*). *Werre* has its origins in Old High German, meaning “confusion, discord, strife”, and goes back to a verbal form in the Old High German, meaning “to bring into confusion or discord”. The OED states that, although the origins of the modern word go back to High German, “no Germanic nation in early historic times had in living use any word properly meaning ‘war’”, that is, meaning *war* in our modern sense of the term. Both noun and verb forms are listed in the entry for *war*. In its entry for the verbal form, the earliest citation is to the Anglo-Saxon Chronicle (dated at 1154). The OED describes this verb as transitive, but notes that this usage is now obsolete. A fuller discussion of the grammatical concept “transitive” can be found in Chap. 5; but for the purposes of this discussion, a clause is transitive when the action construed by the verbal group extends to another participant. In other words, a transitive process has an impact experienced by a distinct entity. By contrast, an intransitive process unfolds without causing a ripple over and beyond the single inherent participant. To exemplify this distinction, consider the two following examples, taken from the OED, which show Middle English versions of *war* as verb in both a transitive and intransitive configuration:

1154 Anglo-Saxon Chron. (Laud) ann. 1135 Dauid king of Scotland toc to uuerrien him.
c1400 Mandeville’s Trav. (1839) xxiii. 251 And whan thei werren, thei werren fulle wisely.

In the example from the Anglo-Saxon Chronicle, the King of Scotland is engaged in an action that extends to and impacts another entity. By comparison, in the example from Mandeville’s Travels, the action of warring has no object. It is an action that does not impact anything. This grammatical distinction between the transitive and intransitive forms is shown in Table 4.1, with Halliday’s terms Actor and Goal

---

4For further details of the corpus composition, see natcorp.ox.ac.uk
describing the functions of the “source of the energy bringing about the change” (Actor) and the entity on which or whom “the outcome is registered” (Goal) (Halliday and Matthiessen 2014, 225–26).

In its earlier life, it appears, war could be used as a verb, and could take a direct object. The meaning of this Old English verb form is given, for the modern English reader, as “to make war upon”, the post-verbal preposition upon reminding us that, in modern English, war, despite the destructive power that is one of its defining characteristics, cannot now construe action in transitive terms. The OED lists the modern word as noun only. Empirically, this can be confirmed by a search of the Google Books corpus, a corpus which contains over 360 billion words of English (Michel et al. 2011). This corpus enables searching by part-of-speech, and returns no instances of the lexical item war functioning as verb or process. War, it seems, has been subject to the slow movement of English towards more and more nominalisation (see Chap. 3).

The OED defines the modern English noun war as:

Hostile contention by means of armed forces, carried on between nations, states, or rulers, or between parties in the same nation or state; the employment of armed forces against a foreign power, or against an opposing party in the state.

Based on the OED citations, this meaning is not quite 900 years old. The earliest use of the word in its current sense is an entry in the Anglo-Saxon Chronicle, dated to 1122:

a1122 Anglo-Saxon Chron. (Laud) ann. 1116 Se cyng Henri fylste his nefan.þe þa wyrre hæfde togeanes his hlaforde þam cynge of France.

[Translation: The aforementioned king Henry assisted his nephew who then had war against his lord the king of France].

As a noun, war is both a countable noun and a mass noun. As a countable noun, war can have a “particularized sense” where it means “A contest between armed forces carried on in a campaign or series of campaigns”. Note that the use of the word to mean “Actual fighting, battle; a battle, engagement” is now obsolete. So while the word can and is frequently used with reference to “particularised” events, these events must be of a certain scale to attract the designation of war. As the OED notes, it is frequently used in this particularized sense with the definite article “to designate a particular war” and “with (an) identifying word or phrase, as in the Trojan war, the Punic Wars, the Wars of the Roses”. The category of war attracts particularizing

---

5 Note that Google’s ngram viewer requires a minimum of 40 instances for a result to be returned.
6 With thanks to Associate Professor Rosemary Huisman for this translation.
names, a feature important for developing narratives around specific wars. War can also be used metaphorically or rhetorically for “any kind of active hostility or contention between living beings, or of conflict between opposing forces or principles”. The dictionary also notes that one can declare, wage, levy or make war.

The OED shows that the modern definition of war entails the notion of “nation” or “state”. Thus, the antagonists to war are political entities, either representing a political body as a whole, or parties internal to a political body. War is also defined as action “between” political parties. Entailed in its definition is action that is bidirectional. This feature of the definition is echoed in an analogy Clausewitz drew on to get at the meaning of war:

we shall do best by supposing to ourselves two wrestlers. Each strives by physical force to compel the other to submit to his will: each endeavours to throw his adversary, and thus render him incapable of further resistance (von Clausewitz 1873, Chapter 1, Section 4.1).

War is, therefore, associated with parties acting on each other. This bidirectionality of action is likely to exclude the semantics of transitive action, making war a form of happening rather than action that extends to, and impacts, outside of itself. We see below that war tends to collocate with middle voice processes: war is construed as beginning, continuing, escalating, ending. In these configurations, war has a life of its own, going through cycles of unfolding without reference to any force other than itself.

Furthermore, both the general and particularized sense of war make reference to “armed forces”, defined in the OED as a body of armed men constituting “the fighting strength of a kingdom or a commander in a field”. The definition of war includes its association with highly organised polities, echoing Malešević discussion of war as organised violence. The action of war is defined as armed forces engaging in “hostile contention” on behalf of a political entity. To try to get more concrete in our definition, let us consult the verbal form to contend (rather than the nominalization contention), which the OED defines as:

To strive in opposition; to engage in conflict or strife; to fight. Const. with, against (an opponent), for, about (an object).

Note first of all that all of the synonyms for contend – strive, engage (in conflict or strike), fight – are intransitive verbs, that is, verbs which construe actions without impact. Fight can involve a direct object, but the potential transitive nature of fight is not made visible here. Instead, a set of typical prepositional associations are made – to fight with, against, for, and about. The latter two, like the verb to strive, construe for us the idea of war as means to some better ends. War is goal-directed, and purposeful. To fight with invokes collective action; and to fight against, oriented to specifying the second party to the conflict, returns us to the bidirectionality that inheres in the semantics of war. Evidence that there is an intransitive semantics associated with fighting can be found in the much greater tendency for the collocation of fought with and died than with and killed. The Google Books corpus gives us an

---

“Const.” is an OED abbreviation meaning “construction” or “construe with”
empirical reading on the much higher probability of *fought and died* over *fought and killed* – see Fig. 4.1.\(^8\) Being a diachronic corpus (one that can show us change over time), Google Books also reveals that the construction *fought and died* recapitulates the war history of the twentieth century, with peaks for *fought and died* in 1919 and 1941. By contrast, *fought and killed* shows no such reaction, suggesting that this configuration is not associated with war.

The OED dictionary entry includes a host of compound words in which *war* is a key component. There are so many that the OED attempts some categorisation of these compounds. The first category is defined as “simple attributive use”, with the sense of “of or belonging to war”, “used or occurring in war”, or “suited or adapted for war”. This category is further sub-categorised as “general use” (40 compounds), words denoting “arms, accoutrement, implements” (23 compounds), words denoting a “commander, officer, army” (7 compounds), words denoting “cries, songs, musical instruments” (11 compounds), words referring to finance (13 compounds), and words denoting “literary or artistic works” (17 compounds). The second category, described by the OED as “objective, etc” (meaning a compound word in which *war* functions as a kind of object), has 18 compounds. The third category includes words that construe war as the instrument of some process, and this category lists 27 compound forms. The final category, described as “special combination”, might perhaps also be described as “not elsewhere listed”: it lists a diverse array of 87 further compound words. *War* has nearly 250 compound forms.

*War*, therefore, acts as Classifier to an enormous variety of things, including family life (*war-bride, war wedding, war baby, war widow, war-orphaned*), artistic forms and types of artist (*war photograph, war poet*), instruments of war (*war-axe, war-boat*), forms of attitude towards war or mental strain/illness associated with war (*war-minded, war strain, war psychosis, war-weary, war fever, war-loving, war hysteria, war-guilt*), social institutions pertaining to war (*War Department, war hospital, war cemetery, war college, war memorial*), the economy or finance associ-

\(^8\)At the peak of 1941, *fought and died* is 0.000022% and *fought and killed* is 0.000002 a ratio of 11–1. At 2000, this ratio is just over 5–1.
ated with war (war economy, war chest, war expenditure, war debt, war bond, war surplus), the effort and achievement particular to war (war production, war footing, war effort, war service, war record, war purpose, war hero, war heroine), space (war zone), time (war years, war period, war generation), and war-related governance (war measure, war-machine, war ration). In these compounds, the premodifier war creates a hyponym. In other words, each compound is a more delicate instance of a general category. Thus, war hysteria is a more delicate form of the category hysteria; a war bride is a particular type of bride.

We also find compounds that construe a reduced form of a transitive relation, in which we find war either as the Actor in the process (“the source of energy bringing about the change”), or as its Goal (the entity on which “the outcome is registered”). In the compound adjectives, such as war-battered, war-ravaged, war-scarred, war-torn, war-weathered, war-wracked, war is the Actor of the processes battering, ravaging, scaring, tearing, wearying, wracking. As a structure, these kinds of compound adjectives give war considerable destructive power, at the same time that they allow the belligerents to be absent from the processes of battering, ravaging, scaring, tearing, wearying, wracking, etc. In other words, being battered or torn by war is construed as a static characteristic of a location or people, rather than an action being done by some people against others. In compounds such as war-breeder, war-chronicler, war-maker, war is the Goal in a transitive clause.

Since nothing, apart from the relationship of signifier and signified, is arbitrary in language, then we need to wonder at this prolific set of compound forms, which can only have evolved through usage, i.e. via parole. We can interpret this proliferation of terms as a measure of the kind and intensity of communal pressure that has been exerted on this lexical item, a process going back nearly a millennium, if we look only from the perspective of our modern English term. But as the quote from Ecclesiastes at the beginning of this chapter indicates, the concept, manifest in various languages, has a long and deep history. The “time for war” of Ecclesiastes – the Hebrew word is milhamah, translated as both war and battle – gives us a collocation with a history well over 2000 years old. With such a long and deep history, it is not surprising that English users have an extensive range of compound forms at our disposal. These compounds indicate that war has considerable reach in the life of English-speaking communities. Moreover, that it is an event requiring organisation is formalised in various ways, such as the names for social institutions which pertain to war, and in the titles that are given for those with degrees of authority in managing the prosecution of war. In addition, we give names to particular types of art forms associated with war. War is not only a naturalised part of English-speaking cultures, simply part of life: it is recorded and memorialised in various art forms. It is not only institutionalised, but culturally absorbed and infused.

One significant implication of these many compound forms is that the word itself must be neutralised. This quality of the semantics of war is important for exploring Malešević’s claims about war being construed as rational, but not emotional. If war can equally appear in the expression war hero and war crime, war itself functions as a taxonomic Classifier, locating some entity with respect to a particular class. In being able to form taxonomic relations, war cannot be a subjective category. War
has to be a neutral lexical item, able to take positive or negative prosodies from its collocations, but bestowing only classificatory semantics itself.

A further resource in developing our picture of the meaning of war is the thesaurus. A thesaurus offers us a semantic classification of a word, in relation to a map of the semantic fields of a language. In a thesaurus, a word is given a location by virtue of its semantic proximity to other words. A thesaurus reverses the principle of organisation of a dictionary, by putting meanings first, and relating meanings to words. It is of interest in our examination of the lexical items war and violence to consider how they are located in a thesaurus, such as Roget’s or the OED thesaurus, and to consider how proximal these lexical items are. Roget’s classification of English lexis begins with six categories of meaning: ABSTRACT RELATIONS, SPACE, MATTER, INTELLECT, VOLITION, and SENTIENT AND MORAL POWERS, a categorisation Roget described as “commencing with ideas expressing abstract relations” and “proceed(ing) to those which related to space and to the phenomena of the material world, and lastly to those in which the mind is concerned” (Roget 1998, xviii). War is located within the category of ideas “derived from the exercise of VOLITION; embracing the phenomena and results of our Voluntary and Active Powers” (Roget 1998, xix). Here we find an association between war and the exercise of rational, measured judgement: war is considered, rational action, rather than emotional response. Under the category of “individual volition”, war is allocated to the subcategory of “antagonism” – see Fig. 4.2.

In the index of the thesaurus, war is also associated with the following entries: destroyer (168), slaughter (362), and dissension (709), the last of these being an entry close to the main entry for war. As we will see below, the entry destroyer is close to the entry for violence in this schema (both are taxonomised as part of the category “causation”, a subcategory within the domain of ABSTRACT RELATIONS). This point is significant when we later consider the potential for the words war and violence to be found in each other’s neighbourhood in the available large corpora. Slaughter also seems a rational choice as a field of semantics with proximity to war. Slaughter is found under “killing: destruction of life” (within the general domain of MATTER, and more specifically “organic matter”). Roget associates slaughter with bloodshed, butchery, carnage, whole murder, massacre, 

Fig. 4.2 War in Roget’s thesaurus
holocaust, extermination, annihilation, among other lexical items. Finally, I note the proximity of war to peace and pacification in Roget’s classification.

The OED also includes a thesaurus, which draws a slightly different initial separation between THE EXTERNAL WORLD, THE MIND, and SOCIETY. War is located within the domain of SOCIETY, which includes the following elements: “the community”, “inhabiting or dwelling”, “armed hostility”, “authority”, “morality”, “education”, “religion”, “communication”, “travel”, “occupation”, and “leisure”. War, not surprisingly, falls into the subcategory of “armed hostility”, but what is noteworthy is that “armed hostility” is construed as part of society. It is of the same order as “morality”, “religion”, “inhabiting or dwelling”, “education”, etc. – see Fig. 4.3. The location of war in this fashion suggests that lexicographers working with one of the most reputable dictionaries in the English-speaking world, and drawing only on actual usage as their evidence, treat war as part of society as much as are the categories alongside which it is collocated.

4.3 Violence in the Dictionary and Thesaurus

The dictionary entry for the word violence is considerably shorter than that for the word war. The OED defines violence as:

The deliberate exercise of physical force against a person, property, etc.; physically violent behaviour or treatment; (Law) the unlawful exercise of physical force, intimidation by the exhibition of such force.

The English word derives from an Anglo-Norman and Middle French term. The earliest citation in the OED is attributed to St. Thomas Beckett in 1300. While war
is a use of force, violence is the “unlawful exercise of physical force”. The word violence has produced a standard set of morphological variants such as violent, and violently (as well as a couple of obscure other forms, e.g. violency, violated). Like war, it has some history in a verbal form. Both violent and violence are listed as verbal forms, probably derived from the Middle French form violenter. Both are transitive constructions, meaning to “compel, coerce, or constrain”, or “to subject (a person) to violence; to attack, assault”. Both are now considered obsolete or rare. Like war, violence is nominal only. Unlike the proliferation of compound forms involving war, according to the OED there are no compound forms based on violence.

While war has nearly 250 compound forms, violence has none. The OED suggests the collocations of domestic, gang and mob with violence, but it does not formally list these as compound forms. Why does violence lack the linguistic infrastructure of war? One part of the explanation may be the fact that violence is inherently illegitimate. This could explain why it does not get used as a reference point for various forms of human endeavours. Moreover, though nominal like war, it does not have the capacity as war does to be countable. This means it can never be used as part of a proper noun. Violence cannot, therefore, stand as a label for a set of actions over a particular time, in a particular location, between particular entities. It lacks the narrativising potential of war.

With respect to the thesaurus, Roget gives violence a distinct location from war (see Fig. 4.4). Violence is located within ABSTRACT RELATIONS, in the subcategory “causation”. At the same time, the index entry sends the reader towards a number of related entries. These include two entries close to the main entry, destruction (165) and vigorousness (174). As noted, Roget relates violence and war. As such, the entry for war (718) is suggested in the index entry for violence, as are two other entries from the domain of VOLITION, misuse (675) and brute force (735). Two further entries are related to violence, and these appear in the last of the six domains of Roget’s classification, that of SENTIENT AND MORAL POWERS. In this category, violence is related to the entries excitable state (822) and cruel act.
attributing to violence emotional associations both intense and malevolent. No such association is made for the term war.

In the OED thesaurus (see Fig. 4.5), violence is located within the general category of THE EXTERNAL WORLD, and, therefore, outside of SOCIETY. The OED’s EXTERNAL WORLD category is further divided into the following subcategories: “the universe”, “the earth”, “the living world”, “sensation”, “matter”, “abstract properties”, “relative properties”, and “the supernatural”. The category “abstract properties”, contains various subcategories (“existence”, “creation”, “cause”, “present events”, “action or operation”, “time”, “space or extent”, “motion or movement”). Of these, the category of “action or operation”, is subdivided into various types, one of which is “behaviour or conduct”. Once again, this category is further subdivided, and includes the category of “bad behaviour”. It is in this category that we finally locate “violent behaviour”. As we can see from its thesaurus location, violence is not a thing in itself, but is a type of action: that is, it is defined as a kind of behaviour or conduct inherently “bad”. Violence is, therefore, a category of action, and to label an action as violence unequivocally delegitimises those actions. This feature of the nominal element violence gives it a distinct relationship to the manifestations of the category from the word war. Violence functions in a hyponymic relation to actions which can attract this label. Hyponymy is a relation of a lexical item to a more general category which contains it. Thus, violence is a higher order category, and by labelling actions as violence, they are attributed with the negative associations of this higher order category. By contrast, the relation between war and its manifestations is one of “meronymy” or part-whole relations. Therefore, events such as invasions and bombings, when construed under the umbrella of war, are given the status of being parts of some larger whole. War func-
tions as an explanatory abstraction which gives purpose to, and in a sense rationalises, the various manifest actions and things with which it is associated. This microgrammatical contrast between war and violence is as profound as the distinctions in atomic structures of, say, gold versus uranium.

### 4.4 War and Violence: Word Frequencies and Dispersion

I suggested in Chap. 1 that part of the legitimacy of war would lie in the lexical item war being both highly frequent and well-dispersed. This claim is partly based on a concept from information theory, referred to as “surprisal”, coined to capture the intuition “that linguistic expressions that are highly predictable, in a given context, convey less information than those which are surprising” (Crocker et al. 2016). Linguistic events with low probability “convey more information than those with high probability” (Crocker et al. 2016). High frequency bestows naturalness on a lexical item. Evidence for the hypothesis about the frequency of an item can now be relatively easily obtained. In the British National Corpus, for instance, there are 27,217 instances of the word war, which, for purposes of comparison, we can express as a normalised frequency of 276.84 per million words (pmw) – see Table 4.2. This frequency varies by register (see Table 4.3): it is higher for the written component of the BNC (289.49 pmw compared with 170.03 pmw in the spoken component). Within the written text types, war peaks in “non-academic prose and biography” (434.18 pmw), and is at its lowest in the “demographically sampled” speech category. It is five times more frequent in “context-governed” speech than the demographically sampled spoken discourse.

But we need to calibrate these frequencies to interpret their significance. We can contrast them, first of all, with the frequency of violence. Violence has a raw frequency of 5507 in the BNC: its normalised frequency is 56.01 per million words. The normalised frequency of the item across the text types of the corpus ranges from 90.84 in newspapers, to 4.02 in the demographically sample speech. War is thus nearly five times as frequent as the word violence, based on the data in the BNC (see Table 4.2). Table 4.4 provides some further data on this topic, comparing normalised frequencies for these terms from a selection of other corpora.

<table>
<thead>
<tr>
<th>Table 4.2</th>
<th>Frequency and distribution of war and violence in the British National Corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of hits</td>
</tr>
<tr>
<td>War</td>
<td>27,217</td>
</tr>
<tr>
<td>Violence</td>
<td>5507</td>
</tr>
</tbody>
</table>

---

9 “Normalised frequency” is a standard measure in corpus linguistics, which allows comparison across corpora of different sizes. A typical NF is a frequency per million words, which I have adopted here.

annabelle.lukin@mq.edu.au
### Table 4.3  Dispersal of war and violence in the British National Corpus, measured by normalised frequency and text dispersion across text type distinctions

<table>
<thead>
<tr>
<th></th>
<th>War</th>
<th>Violence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normalised frequency</td>
<td>Text dispersal</td>
</tr>
<tr>
<td></td>
<td>Text dispersal</td>
<td>Normalised frequency</td>
</tr>
<tr>
<td>Written</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All written</td>
<td>289.49</td>
<td>1959/3140 = 62%</td>
</tr>
<tr>
<td>Academic prose</td>
<td>287.55</td>
<td>251/497 = 50%</td>
</tr>
<tr>
<td>Fiction and verse</td>
<td>157.58</td>
<td>332/452 = 73%</td>
</tr>
<tr>
<td>Non-academic prose and biography</td>
<td>434.18</td>
<td>530/744 = 71%</td>
</tr>
<tr>
<td>Newspapers</td>
<td>251.8</td>
<td>293/486 = 60%</td>
</tr>
<tr>
<td>Other published written material</td>
<td>260.21</td>
<td>446/710 = 63%</td>
</tr>
<tr>
<td>Unpublished written material</td>
<td>186.72</td>
<td>107/251 = 43%</td>
</tr>
<tr>
<td>Spoken</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All spoken</td>
<td>170.03</td>
<td>245/908 = 27%</td>
</tr>
<tr>
<td>Demographically sampled</td>
<td>46.76</td>
<td>57/153 = 37%</td>
</tr>
<tr>
<td>Context governed</td>
<td>254.54</td>
<td>188/755 = 25%</td>
</tr>
</tbody>
</table>

### Table 4.4  Relative (normalised) frequencies of war and violence in a variety of other corpora

<table>
<thead>
<tr>
<th>Corpus name</th>
<th>Date</th>
<th>Size</th>
<th>war</th>
<th>violence</th>
<th>ratio war/violence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lampeter Corpus</td>
<td>1640–1740</td>
<td>1,135,231 words</td>
<td>359.3*</td>
<td>60.78</td>
<td>5.9</td>
</tr>
<tr>
<td>British Hansard</td>
<td>1803–2005</td>
<td>1.6 billion</td>
<td>582.31</td>
<td>46.27</td>
<td>12.47</td>
</tr>
<tr>
<td>Time Magazine</td>
<td>1923–2006</td>
<td>100 million</td>
<td>882.42</td>
<td>75.5</td>
<td>11.69</td>
</tr>
</tbody>
</table>

* Figure includes “war/warr/warre”
Lampeter Corpus, a collection of early Modern English tracts and pamphlets, shows a similar relative frequency of war to violence. The British Hansard Corpus, 1.6 million words of Hansard transcription from the British parliament from 1803 to 2005, shows a much higher relative frequency (12.47), and a normalised frequency for war more than twice that of the BNC. The Time Magazine corpus, of 100 million words spanning nearly 90 years, shows an even greater preoccupation with war (NF = 882.42). I have also included the figures from the Corpus of Contemporary American English. Notably, the American corpus has a higher normalised frequency for both war and violence, compared with the British National Corpus, but a lower relative frequency for these two items.

The figures in Table 4.4 give us a wider, and historically deeper, basis for the frequency figures based on the BNC. But at the frequency of 276 pmw, we need to wonder whether this figure means that the word war is frequent in an overall sense. Can we also get a frequency ranking for this term? Leech et al. (Leech et al. 2014) produced frequency lists for English based on the BNC, and using raw lexical items. From the perspective of this large, multigeneric corpus, war ranks at number 323 (see http://ucrel.lancs.ac.uk/bncfreq/lists/ 1_2_all_freq.txt), and has a frequency comparable even to some closed system items (e.g. since, towards, anything). It ranks above lexical items such as mother, education, and work. A further source for calibrating overall frequency of this lexical item is Brezina and Gablasova’s “New General Service List” (New-GSL), a core vocabulary word list for English based on four language corpora totalling 12 billion words (Brezina and Gablasova 2013). The list, which is based on lemmatized items not raw lexical items, puts war in the top 500 most frequent words of English, coming in at number 438 in this list. At a similar ranking to war, we find words as common as common (434) and boy (435). War turns out to have a higher ranking than die (464), food (468), father (477), human (504) and love (574). Violence also turns up on the New-GSL, but at number 2107.

With regard to dispersal, high frequency words are logically likely to be dispersed across a number of registers. In the BNC, war is found in over half the texts of the whole corpus (2204 out of a total of 4048). Per text in which the lexical item appears, the average usage is 121 times per text. Text dispersal over the domains of

---

10 The Lampeter Corpus consists of a decade by decade set of texts collected under the following categories: religion, politics, economy, science, law, and miscellaneous. See https://www.tu-chemnitz.de/phil/english/sections/linguist/real/independent/lampeter/ lamphome.htm.

11 The Corpus of Contemporary American English consists of texts from various “genres” of spoken discourse (transcripts of unscripted conversation from radio and TV), fiction, popular magazines, newspapers and academic journals. Although it contains spoken discourse, it lacks any discourse from natural casual conversation.

12 Brezina and Gablasova used three principles of selection for this list: frequency, dispersion, and distribution across language corpora. These measures, they argue, “guarantee that the words selected for the new vocabulary list are frequently used in a large number of texts and that the wordlist is compiled in a transparent and replicable way” (Brezina and Gablasova 2013, 3).

13 The use of both whole texts and of text extracts in the BNC makes it problematic to draw firm conclusions about the dispersal of a lexical item across distinct contexts of situation.

annabelle.lukin@mq.edu.au
the corpus ranges from 73% of all texts in the “fiction and verse” category to 25% of all texts in the “context-governed” spoken discourse. High frequency and dispersed usage mean that war carries little or no evaluative prosody – it has low surprisal value. It is as if no subjectivity or bias is exercised in the choice of this lexical item. The frequency and textual spread of war in the BNC (as well as the other corpora on which Brezina and Gablasova based their core vocabulary list) also gives further evidence to my earlier argument which I based on the proliferation of compound forms: war is textured into many and varied domains of ordinary life. This has to have the semantic effect of naturalising and normalising war. This registerial distribution also gives evidence for Bourke’s claim that “military practice, technology and symbols have invaded our everyday lives” (Bourke 2015, 3). The dispersal of violence is, not surprisingly, lower than for war. The word appears in 1087 texts, or 26% of texts in the BNC. In the texts in which the word appears, its average usage per text is 5. The range of its text dispersion is from 44% of the “fiction and verse category” (also the most dominant category for the use of war, measured by text dispersion), to 6% of the demographically sampled speech. Comparing the two lexical items, the overall relative frequency of war to violence in the BNC is nearly 5:1. The hit per text ratio of war to violence is 24:1. With its lower frequency and small scale of contexts of operation, violence has a higher “surprisal” value than war.

We have, then, a pair of terms, which potentially go together, but which display quite different patterns of frequency and dispersion. One is highly frequent, highly dispersed, experientially neutral by having been used in so many different contexts, and lacking any associated subjective qualities. War, for instance, would not attract any “appraisal” value in the terms set out by Martin and White (Martin and White 2005). The other is less frequent, less dispersed – so more specialised – inherently negative, and carries a higher information value.

4.5 War and Violence: Collocational Patterns

Let us now turn to consider the collocational patterns of these two lexical items, to see what collocation patterns show us about the meanings associated with war and violence. Collocation, Firth argues, is not “mere juxtaposition”, but is rather “an order of mutual expectancy”: words are mutually expectant and mutually prehended (Firth 1957a, 12); and they indicate a “multiplicity of systems derived from carefully contextualized structures” (Firth 1968b, 18). Firth’s conception is widely cited in corpus linguistics. The advent of sizeable, searchable corpora has opened up

14 However, Bartsch and Evert argue in a recent paper that there have been no studies based on a Firthian view: “there has not, to our knowledge, been any systematic large-scale study resting on a Firthian notion of collocation. Studies typically take as their vantage point specific types of multi-word expressions (such as support verb constructions or verb-particle constructions...) or rely completely on intuitions of annotators (e.g. lexicographers’ judgements)” (Bartsch and Evert 2014, 50).
considerable opportunities to advance both linguistic theory and method in relation
to habitual patterning in lexicogrammar, as well as the pressure to operationalize the
notion of collocation for empirical research (Bartsch and Evert 2014, 48–49). Bartsch and Evert define collocation as “the habitual and recurrent juxtaposition of
words with particular other words” (Bartsch and Evert 2014, 48). Baker et al. (Baker et al. 2008, 278) include the concept of a “span” in their definition of collocation,
which is “the above-chance frequent co-occurrence of two words within a pre-
determined span”. Baker et al. note that collocates do not need to be adjacent to the
node to influence the meaning of node word: five words to left and right of the node
word is a typical “collocation window” for the investigation of collocational pat-
terns (Baker et al. 2008, 278). Such a span gives a way of bringing out the “more
general associations” of a particular item under investigation (Brezina et al. 2015).
All collocations discussed in this book are based on a span of five words to left and
right of the node word.

The methods for exploring, measuring and interpreting collocations are the sub-
ject of considerable debate amongst scholars (see e.g. Bartsch and Evert 2014;
Brezina et al. 2015). Three basic criteria relevant to the study of collocation are: (1)
distance from the node word; (2) frequency of the collocational pattern; and (3)
exclusivity, in other words, the degree to which the collocation relation is exclusive
such that two selected words more typically go together than with other words
(Brezina et al. 2015). Brezina et al. (2015) note as an example the expression “in
love”: while love attracts in, in has no particular affinity with love. In other words,
the relationship between in and love “is not exclusive” (Brezina et al. 2015, 140).
Other criteria for measuring collocation strength include dispersion (the spread of
the node and its collocations across a given corpus), and type-token distribution (the
“level of competition” for a slot around a node word from other collocate types) (see
e.g. Gries 2013). Brezina et al. (2015) add a further feature of collocational patterns,
which they describe as the “connectivity between individual collocates” (Brezina
et al. 2015, 141). They note that collocates of words “do not occur in isolation, but
are part of a complex network of semantic relationships which ultimately reveals
their meaning and the semantic structure of a text or corpus” (Brezina et al. 2015,
141).

Various statistical measures of collocational strength have been developed in
corpus linguistics. Establishing an association measure requires a measure of the
“random co-occurrence baseline” (named the “expected frequency”, where
“expected” here means expected by a computer and not a human), compared with
the observed collocational frequencies. The various association measures “can be
understood as different ways of comparing the observed and expected values, put-
ting different weight on different aspects of the collocation relationship” (Brezina
et al. 2015, 145). The collocational relationship is complex, so that “no single asso-
ciation measure can capture all of its aspects” (Brezina et al. 2015, 144). All are
concerned with testing the hypothesis that the co-occurrence – the collocation – is

---

15 How collocational graphs demonstrate “semantic structure” is not made entirely clear by Brezina et al. 2015.
These measures include mutual information (MI, as well as MI2, and MI3), log likelihood, dice co-efficient, T-score and Z-score. The British National Corpus makes a set of these calculations available, including mutual information – a measure of “mutual attraction” – which I will use to retrieve collocations. In particular, I will use the MI3 value, since, as Daille (1995) has argued, it provides a balance between exclusion and inclusion of lower frequency collocations. Thus, MI3 gives more weight to frequency; and, as Brezina et al. (2015, 160) argue, “no matter how suggestive, an association which is not repeated enough will be less influential than an association that is more firmly established in the discourse“. Mutual information is a calculation of the relation between observed collocate frequency and “expected” collocate frequency (where “expected” is determined by the relative frequencies of each of the two items under consideration). The higher the value, the more frequent the association between the node word and its collocate. In line with the recommendation in Brezina et al. (2015) that a standard notation be developed for making explicit the parameters used in extracting collocations for a given node word from a given corpus, the “collocation parameters notation” (CPN) for the collocations of the node words war and violence using the whole BNC is set out in Table 4.5.

Table 4.5  Collocation parameters notation for war and violence in the British National Corpus

<table>
<thead>
<tr>
<th>Statistic name</th>
<th>Statistic cut-off value</th>
<th>L and R span</th>
<th>Minimum collocate frequency (C)</th>
<th>Minimum collocation frequency</th>
<th>Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>MI3</td>
<td>9</td>
<td>5L-5R</td>
<td>5</td>
<td>5</td>
<td>N/A</td>
</tr>
</tbody>
</table>

With the collocation frequency set at 5, and an MI3 threshold value of 9, the BNC returns 1202 collocates for war, and 453 for violence. War has 2.65 times the number of collocates of violence, providing yet further evidence for the argument that war is a concept with wide reach in the lives of English speakers. With so many collocates for both these words, it is not feasible to explore them all. Appendix 1 sets out the top 100 collocates for war and violence in the BNC. The cut-off of 100 is arbitrary, and is simply a means to acquire a working list. Even at 100, the MI3 score is relatively high (>17 for war, and >12 for violence), suggesting that there are many further significant collocational relations with the word war.

Each of the two node words prehend themselves. For war, the explanation is in the point made earlier: the word has so many lexical options that war is found in the environment of war. As a consequence, when war becomes topical, there are many structural forms into which it can enter. There is, by contrast, a singular explanation for why violence prehends violence: it is the oft-heard expression that “violence

16 I note Kilgariff’s claim that “the problem for empirical linguistics is that language is not random, so the null hypothesis is never true”. However, “probability models have been responsible for a large share of progress in the field in the last decade and a half. The randomness assumptions are always untrue, but this does not preclude them from being useful” (Kilgariff 2005, 264).

17 This MI3 value is used as it is the default value for MI3 used in GraphColl (Brezina et al. 2015).

18 For this list, collocates of punctuation were removed (including enclitic ‘s), as well as some function words (e.g. conjunctions, deictics).
begets violence” – see Fig. 4.6. Recalling this common aphorism, we see a rhetorical contrast between war and violence. While our collective wisdom tells us that violence leads to more violence, the “Great War” was thought, by some at least, to be the war that would end wars (Wells 1914). Below, I discuss the legal veneer of the word war, and its long association with the word peace.

The two node words also exhibit some collocational attraction, measured by an MI3 score of 11 (see Table 4.6). There are a total of 32 instances where these terms are collocates within the parameters set for this study – recall that Roget’s thesaurus entry also suggested a relationship between these words. The collocations of war and violence are dispersed across 29 texts. Since war appears in 2204 texts in the BNC, this means that in just over 1% (1.3%) of the texts in which war appears the word violence is also present (n = 1087). The figure is 2.7% for the presence of war in texts that feature the word violence. While speakers of English are free to put these words together, the evidence from the BNC is that collectively we maintain a strong separation between them.

Corpus linguists analyse collocations by seeking “semantic prosodies”. McEnery (2006, 20) uses the term “semantic prosody” to refer to the “meaning arising from the interaction between a given node word and its collocates”. This is a Saussurean principle: meaning arises in the play of signs, and so syntagmatic relations are key to the meaning of a word, both in an instantial and in a systemic sense. McEnery also argues that semantic prosodies typically encode attitudes and evaluations, and he goes on to suggest that the evaluations conveyed by such prosodies are typically negative (although he does not explain or elaborate this claim) (McEnery 2006, 23). Semantic prosodies are more than simply positive versus negative, as McEnery’s discussions of the collocations and keywords of his corpora illustrate.

Fig. 4.6 “Violence begets violence”: BNC concordance lines for violence and begets

Table 4.6 Collocation of war and violence in British National Corpus

<table>
<thead>
<tr>
<th>Node word</th>
<th>Ranking</th>
<th>n</th>
<th>n expected</th>
<th>n observed</th>
<th>n texts</th>
<th>Coll. value</th>
</tr>
</thead>
<tbody>
<tr>
<td>War</td>
<td>Ranked at 239 in collocations for violence</td>
<td>27,217</td>
<td>11.523</td>
<td>32</td>
<td>29</td>
<td>MI3:11.4735</td>
</tr>
<tr>
<td>Violence</td>
<td>Ranked at 596 in collocations for war</td>
<td>5507</td>
<td>11.509</td>
<td>32</td>
<td>29</td>
<td>MI3: 11.4754</td>
</tr>
</tbody>
</table>
Table 4.7 Semantic prosodies: collocational patterns in the top 100 collocations of *war* and *violence*

<table>
<thead>
<tr>
<th>Premodifier of nominal group structure</th>
<th>War</th>
<th>Violence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Classifiers</strong></td>
<td>Geographic: world, Gulf, Boer, Vietnam, Falklands, Korean, Crimean, Iran-Iraq, Franco-Prussian, Spanish, Iraq, Peloponnesian, American, Imperial, Russo-Japanese, Arab-Israeli, Pacific</td>
<td>Geographic: natal</td>
</tr>
<tr>
<td></td>
<td>Ordinal: second, first, II, last</td>
<td>Ordinal:</td>
</tr>
<tr>
<td></td>
<td>State-internal: civil</td>
<td>State-internal: township, townships, sectarian, ethnic, loyalist, communal, intercommunal, mob, PAV/VPN, street</td>
</tr>
<tr>
<td></td>
<td>Means: cold, nuclear</td>
<td>Means: physical, sexual</td>
</tr>
<tr>
<td></td>
<td>Extent: great, all-out</td>
<td>Extent: widespread</td>
</tr>
<tr>
<td></td>
<td>Goal-directed (including lack of):</td>
<td>Goal-directed (including lack of): drug-related, political, anti-foreigner, racial/racist, domestic, mindless, gratuitous</td>
</tr>
<tr>
<td></td>
<td>Gender:</td>
<td>Gender: male</td>
</tr>
<tr>
<td></td>
<td>Legality: just(^b)</td>
<td>Legality: unlawful</td>
</tr>
<tr>
<td>Epithet</td>
<td>Phoney</td>
<td>Sporadic, serious, worst</td>
</tr>
<tr>
<td>Process (including in nominalised form)</td>
<td>Relational/auxiliary</td>
<td>Be/is/was/were/been, has/have/had</td>
</tr>
<tr>
<td></td>
<td>Material and transitive</td>
<td>Brought killed won</td>
</tr>
<tr>
<td></td>
<td>Material and intransitive</td>
<td>Waged/waging, broke (out), end/ended/ending came began/beginning outbreak</td>
</tr>
<tr>
<td></td>
<td>Material and both transitive and intransitive:</td>
<td>Fight/fought/fighting</td>
</tr>
<tr>
<td></td>
<td>Verbal and intransitive:</td>
<td>Declared/declaration,</td>
</tr>
</tbody>
</table>
| | Verbal and transitive: | Threat/threats/threatened/threatening, advocated, condone, abuse | (continued)
Table 4.7 (continued)

<table>
<thead>
<tr>
<th>Common nouns</th>
<th>Associated things (various)</th>
<th>War</th>
<th>Violence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Peace, aftermath, crimes, memorial, time, effort, years, period, attrition, horrors, laws, Museum, Office, wagon, trade, revolution, independence, boars</td>
<td>Sex, crime/s, offences, vandalism, robbery, disorder, bloodshed, fear, cruelty, football, dishonesty, corruption, hatred, looting, campaign, aggression</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Proper nouns</th>
<th>Associated places</th>
<th>France, Germany, Britain, Russia, zone, Europe,</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prisoner/s, veterans, criminals, military, guerrilla, victims, Hitler, hero, army</td>
<td>Women, victims, police, IRA, spectator, Sind</td>
</tr>
</tbody>
</table>

*a ‘VPN’ and ‘PAV’ are the Slovakian and English acronyms for a political movement established in Slovakia in 1989 — see https://en.wikipedia.org/wiki/Public_Against_Violence*

*b ‘Just’ is a multifunctional lexical item: it can operate as modal Adjunct, or as an experiential element, i.e. as Classifier in a nominal group structure. Its polysemy explains its collocational frequency with war (MI3 = >17; n observed = 337). Only a subset of these instances related to the concept of “Just War”

Table 4.7 sets out the top 100 collocates of both war and violence, but organised, in the first instance, by a grammatical principle. The first column of Table 4.7 distinguishes the lexical items into three grammatical categories: premodifiers of war, process items (both as verbal group elements, and where processes have been nominalised), and common and proper nouns. Premodifiers are then split into Classifiers and Epithets; and Classifiers are further subcategorised using the following semantic categories: “geographic”, “ordinal”, “temporal”, “state-internal”, “means”, “extent”, “goal-directed”, “gender”, and “legality”. These categories I derived in an ad hoc manner, based on my sense of semantic proximities between groups of collocates. Verbal group elements are differentiated based on Halliday’s process type schema: “relational” (which are also potentially auxiliary verbs), “material and transitive”, “material and intransitive”, “material and both transitive and intransitive”, “verbal and intransitive”, and “verbal and transitive”. Finally, I make further simple distinctions in the category of “common nouns” between those that refer to humans and human collects, and all other entities (see discussion of “thing type classification” in Halliday and Matthiessen 1999).

What do these associations show? Note first of all there are no positive associations with the word violence. While some associated terms are neutral, many are either always negative (sectarian, racist, unlawful, escalate, rape, flared, denounce, vandalism, robbery, disorder, cruelty, hatred, looting), or they become so when coupled with violence, as in the collocation of domestic, sexual or physical with violence. And not only does violence collocate with unlawful, it also collocates with various acts deemed unlawful in most if not all jurisdictions: vandalism, robbery, looting. The term offences means unlawful acts. Violence also collocates with aggression (the MI3 score is 15.9203; there are 32 instances across 12 texts), a word with an overall much lower frequency than that for either war or violence (the frequency of war to aggression in the BNC is nearly 22/1). With respect to the process
collocations, violence attracts transitive verbs, both material and verbal. As explored earlier, violence itself comes from a transitive verb. Violence is something that has an impact. I note too that, in this comparison, the lexical verbs associated with violence are greater in number than those for war (though we must bear in mind we are dealing only with the top 100 collocations). These items also tend towards being lexis with a reasonable degree of specificity. In terms of the range of Classifiers, violence is classified with respect to forms of state-internal violence (attracting ten collocates in this category), by the means adopted, and by its goal orientation (political, anti-foreigner) or lack of same (mindless). Violence is also gendered, taking the Classifier male.

The prosodies around war are mixed. There are some negative associations (horrors, criminals and crimes); and war, like violence, has victims. But war is also associated with the word hero, with peace, with independence and revolution. It also has many collocations that are non-evaluative – being concerned with time and place. These collocates locate individual instances of the category war, providing the schema for individuating wars and giving them proper names, which is a crucial step in the process of them becoming the topic of historical narrative, and to their being memorialised and commemorated. In this top 100 list we find a preoccupation still with a war which is c. 2500 years old, the Peloponnesian War. That modern British discourse, across 100 million words and various quite distinct contexts, is still preoccupied with the wars of Ancient Greece is perhaps explained by Hanson’s claim that “The Greek and Roman writers who created the discipline of history defined it largely as the study of wars” (Hanson 2010, 3). This historical orientation to the recounting of war no doubt explains the collocations of war with memorials and museums. War is also associated with Office (see Fig. 4.7), a sign of the institu-

Fig. 4.7 Collocation of war and office in the BNC
tionalised nature of war (Malešević 2010). The capitalisation of this lexical item is a sign of the official status: it is a reference to the British War Office, an institution with its origins in the seventeenth century.

From 1964, the “War Office” became the Ministry of “Defence”, a noteworthy resemanticising of this government function. Despite this change, the “Office” remains a dominant collocation even into the 1990s.

Recalling the claim from Malešević that, in the bureaucratisation of coercion, internal and external forms of violence are made distinct, I draw attention to the various lexical items in collocation with violence that construe a state-internal meaning. From the top 100, there are 10 items of this nature that collocate with violence, such as township/s, ethnic, sectarian, mob. By comparison, there is only one for war, and it is the pre-modifier civil. Despite civil war construing state-internal violence, the concept does not have the sense of disorder associated with a concept such as sectarian violence.

While violence is inherently unlawful, war is not. Crime collocates with war, but we also find the collocate laws. War is governed by international law – there are “laws of war”. A “war crime” is a crime against the laws of war. By implication, war is not, per se, illegal, although the problems of determining whether the use of armed force by a state constitutes a sanctioned use of violence appear intractable. Examples of attempts by the United Nations to draw boundaries around terms such as armed conflict, or aggression, put on display the limits of language. Aggression, for instance, is a term within international law meaning the unlawful use of force. Chapter 7 of the UN Charter is titled “Action with Respect to Threats to the Peace, Breaches of the Peace and Acts of Aggression”, and is the mechanism by which the Security Council can authorise “action”, including “military action”, against a state deemed to have breached the charter. The UN did not achieve consensus on the meaning of aggression until 1974, the culmination of over 50 years of discussions through the tenure of both the League and the United Nations (Stone 1977). According to Wilmshurst (2008), the definition has “scarcely ever been used for its primary purpose as a guide to the Security Council in determining aggression by States”. Phrases in this corpus associate aggression with Japan, with Islam, and with Serbia. The clear implication is that these states or religious groups are not legitimate in their use of violence. Aggression, like violence, delegitimises the use and users of force.

Through collocational patterning – the “company that this word keeps”, to reiterate Firth’s (1957a) well-known phrase – war is a lexical item invested with a connotation of legality, and by extension, legitimacy. “Legality” is one of Weber’s three principles by which the exercise of force or power gains legitimacy. Rule through legality is based on a “belief in the validity of legal statutes and practical ‘competence’ based on rationale rules” (Weber 2004, 34). Although war can be associated

---

103 with illegality, it requires premodification for this negative meaning to be invoked, unlike violence which, like a snail carrying its shell, cannot divest itself of its illegitimacy. And not only can war be legal, it can be “just”, a collocation that also appears in the BNC (n observed = 30; see Fig. 4.8), and which imputes a moral dimension to the business of war which is unthinkable with the violence. The concept of “Just War” has a long and deep history in Western culture, such that it is possible to claim that “we in the West cannot think about war without using the terms of this broad tradition, even if we disagree with what it teaches” (Johnson 1986, 2). The precise origins of the philosophical debate are obscure, though it is developed in Aquinas’ late thirteenth century Summa Theologica (Question 40 in The Second Part of the Second Part: On War), in which he asks whether is is always sinful to wage war. Aquinas is in dialogue with Augustine (354–430 AD), who mused on these matters some seven centuries before Aquinas. Both theologians concluded that war was not sinful. Aquinas elaborated the conditions under which

---

20 I draw on the translation published online by The Aquinas Institute (see http://theaquinasinstitute.org/about/), and available online: http://aquinas.cc/
he claimed war was justified: that it be waged under the authority of a sovereign in whom responsibility is vested “to watch over the common weal of the city, kingdom, or province subject to them”; that it be waged for a just cause; and that belligerents should have a “rightful intention”. It is possible to see how this kind of theological musing on war helped create a deep legitimacy of the category. These tracts predate, but influence, the emergence of secular debates on the nature of war, and are part of the history of the development of international laws of war. It is a rather obvious point that a treatise on the legal basis of war not only proscribes certain forms of violence, it validates others. It establishes that war is legitimate, even if conditions must be met for it to be so.

A final observation on the BNC collocates is the association of the word peace with war. Modern English, as represented in the BNC, observes an association of war with peace (MI3=19.1082; n expected=18, n observed=217). Peace is number 32 in the list of collocations of war. The association turns up in 145 of the 2204 texts in which war is found in the BNC, that is, in 6.6% of texts. Based on the BNC, the probability that a text in which the word war is found will show a collocation of war with peace, compared to war with violence, is 5–1. The association of war with peace has an extended “semohistory” (Halliday 2003). Augustine linked peace to war when he argued, “We do not seek peace in order to be at war, but we go to war that we may have peace. Be peaceful therefore, in warring, so that you may vanquish those whom you war against, and bring them to the prosperity of peace” (cited in Aquinas’ Summa Theologica). Aquinas adds in defense of war that “Those who wage war justly aim at peace, and so they are not opposed to peace, except to the evil peace, which Our Lord ‘came not to send upon earth’”. Grotius, the “father” of modern international law, restates this association in the opening of his 1625 treatise On the Rights of War and Peace (De Jure Belli et Pacis): “War ... is undertaken in order to secure peace ... War itself will finally conduct us to peace in its ultimate goal” (Grotius 1901). Article 42 on the UN Charter gives the Security Council the legal basis to “take such action by air, sea, or land forces as may be necessary to maintain or restore international peace and security” – the right, in other words, to use the means of war for peaceful outcomes. Such formulations, over a time span of some 1700 years, show us a deep contradiction in our collective views on the nature of war.

4.6 Conclusion

This chapter has explored the legitimation of war by examining the definition, etymology, frequency and collocational patterns of the lexical item war, by comparison with the lexical item violence. As predicted, war is a highly frequent lexical item, shown in the BNC and new GSL lists to be as common as words such as common and boy, and higher in ranking than food, mother, and work. It is widely dispersed across register types, and has many compound forms – almost 250. In addition, its
collocational patterns show largely neutral (e.g. taxonomic) or positive associations. While some negative associations are possible, this lexical item is not inherently negative. By contrast, violence is a lexical item that is used as a form of negative evaluation. Unlike war, violence is, by definition, a negative category. While violence has a transitive origin and retains this meaning, war in modern usage appears oddly intransitive. Not surprisingly, its collocates are negative. Its much lower frequency and dispersion mean that the lexical item is more likely to carry “surprisal” value. The thesaurus gives these lexical items distinct locations; and the collocation analysis based on the BNC shows that these two words are unlikely to be found in proximity to each other. War is more likely to collocate with peace than with violence, by a factor of five to one. The BNC gives us evidence that these two lexical items represent distinct categories, strongly bounded from each other. War, unlike violence, is also associated with a striving towards some better state. In Chaps. 5 and 6, the separation of these two lexical items is explored by contrasting two news reports on the 2003 invasion of Iraq, one of which construes the invasion as war, the other construing it as violence.

References

Daille B (1995) Combined approach for terminology extraction: lexical statistics and linguistic filtering. Lancaster University, UCREL, Lancaster
Gries ST (2013) 50-something years of work on collocations. What is or should be next. Int J Corpus Linguist 13(4):403–437
Wells HG (1914) The war that will end war. F. & C. Palmer, London