

تحليل مدوّني تعاقبي لمصطلحات الحرب

علي حسين حويل

جامعة ذي قار/كلية التربية للعلوم الانسانية

ali.hussein.huwail@utq.edu.iq

المخلص

تُجري هذه الدراسة تحليلاً لسانيًا تزامنيًا-تعاقبيًا (Diachronic) لمصطلحات الحرب خلال الفترة بين 2005 و2020، بهدف سدّ الفجوة في التحليل اللساني المنهجي والمنظم لكيفية إسهام خطاب الحرب في تشكيل الذاكرة الجمعية والحكم العام. وتسعى الدراسة إلى تتبع التحولات اللغوية في مصطلحات الحرب وقياسها وتحليلها، مع افتراض أن الأنماط اللغوية لا تقتصر على السياقات السياسية فحسب، بل تمتد أيضًا إلى الأبعاد الثقافية والاجتماعية. وبالاعتماد على إطار تحليل مدوّنات لغوية (Corpus) يدمج بين التحليل النقدي للخطاب (CDA) واللسانيات الإدراكية، تستكشف الدراسة التغيرات ذات الدلالة الإحصائية عبر ثلاث مراحل زمنية متميزة. وتشير النتائج الأولية (2005-2011) إلى تركيز على خسائر الجيش الأمريكي وأحداث صراعية محددة (العراق/أفغانستان). أما التحول الاصطلاحي الثاني (2014-2015) فيشمل النزاعات العسكرية، والجهات التنظيمية الفاعلة، والسياقات الدبلوماسية. وأخيرًا، تُظهر الفترة (2017-2020) انخفاضًا حادًا في تكرار مصطلحات الحرب، حيث يُعاد تمثيل الحرب أساسًا من خلال الإحالة إلى أحداث تاريخية (مثل "نهاية الحرب العالمية الثانية") أو سياقات مؤسسية (مثل "فيلق المهندسين في الجيش الأمريكي"). ويشير ذلك إلى احتمال حدوث تلاشي دلالي (Semantic Bleaching) وتحول في الوظيفة نحو استخدام تداولي ذي طابع تاريخي أو تذكاري.

Diachronic Corpus Analysis of War Terminology

Ali Hussain Huwail

College of Education for Humanities, University of Thi-Qar, Thi-Qar, Iraq
ali.hussein.huwail@utq.edu.iq

Abstract

This paper conducts a diachronic linguistic analysis of war terminology between 2005 and 2020 to bridge the gap in methodologically structured, systematic linguistic analysis of how war discourse shapes collective memory and public judgement. The aim is to trace, quantify, and analyze the linguistic shifts in war terminology, hypothesizing that linguistic patterns are not only in political contexts but also in cultural and social perspectives. Following a corpus analytical framework that integrates CDA with cognitive linguistics, this study explores statistically significant changes observed across three distinct periods. Initial findings (2005-2011) show an emphasis on U.S. military casualties and specific conflict events (Iraq/Afghanistan). The second terminological shift (2014-2015) encompasses military conflicts, relevant organizational actors, and diplomatic contexts. Finally, the period (2017-2020) encompasses a remark about the frequency of war terminologies, which are sharply decreased, where war is primarily constructed through reference to historical events ("the end of World War II") or institutional contexts ("U.S. Army Corps of Engineers"). This suggests potential semantic bleaching and a shift in function toward historical or commemorative pragmatic use.

Keywords: Diachronic analysis, corpus linguistics, war terminology, semantic preference, geopolitical dynamics

Introduction

The study of war echoes as a core idea within the domain of strategic analysis and international relations because of its role as the definitive essential mechanism that determines the configuration of the international system and the most destructive means by which states go after their strategic priorities. In the absence of a global order, the possibility of military conflict plays a decisive role in ordering global power, determines who allies with whom, and demands ongoing and persistent research in the field of national security and deterrence planning. Additionally, this field of study remains of critical importance since the nature of armed conflict is ever-changing, necessitating academic analysis of emerging forms of aggression. By examining war, researchers aim not only to explore how states employ military power to fulfil political aims (according to Clausewitz's theory) but also to confront the tragic human suffering and economic devastation, making it a fundamental resource for guiding policy decisions directed toward ensuring the stability of the international system.

This paper addresses a key gap in structured, wide-ranging diachronic linguistic analysis by exploring how the language of war shapes collective memory and judgment or how its mediation of the relationship between politics and violence has been subject to long-term variation. The problem statement focuses on the challenges of evaluating and quantifying the linguistic constructions through which war is conceptualized and rationalized. Traditional approaches of political and historical analyses, to some extent, disregard the linguistic mechanisms such as metaphor and euphemism through which the concept of war is discursively shaped and constructed. For this reason, the principal objective is to conduct a comprehensive corpus analysis to trace, quantify, and analyze these shifts in terminology of war between 2005 and 2020. More specifically, it aims to investigate how the conceptualization of war has changed over time. It assesses how these linguistic changes align with geopolitical contexts and identifies linguistic features that reflect the 'war fatigue.' The key hypothesis is that war terminology is shaped not only by reaction to political contexts but also reveals cultural and social norms, such as the emergence of the war fatigue phenomenon, through systematic shifts or linguistic conventions across time. This research aims to bridge the knowledge gap through systematic empirical analysis connecting linguistic changes with to shifts in societal and political contexts.

This study is critical since it provides a systematic and evidence-based evaluation regarding the progressive changes in war-related language. It provides an insight of how war discourse is associated with public perception and major political significant events. The research, by investigating war terminology diachronically, can indicate transformations in social norms, including rising war fatigue, which shapes and impact levels of public support, and the way it is portrayed in media discourse. Furthermore, examining such transformations in war discourse offers meaningful critical perspectives for researchers and decision-makers who are aware how linguistic patterns mirror and influence political and societal structures. This paper seeks to fill a critical gap within the field of linguistics and geopolitical dynamics by correlating a direct connection between language and socio-historical changes over a critical 15-year period.

Language of War

Considering the critical nature of the topic, the study of war remains a key focus or a principal theme within the field of strategic research and international affairs (e.g., Addison; Jones and Baumgartner 2005; Black 2009; Levy 2014; Brands 2023). The language of war has transformed as quickly in a manner analogous to the conflicts being analyzed. Moreover, war-related terminology, as it is a domain-specific lexicon, reveals shifts in political strategies, technological progress, and ideological or cultural perspectives.

In its essence, the destructive nature of war produces “the death and maiming of human beings, creates widows, [widowers,] and orphans, destroys homes and communities, and wrecks lives.” (Sheffield 2010, 4). Sheffield’s viewpoint draws attention to the impact of war on individuals by redirecting the focus away from power-oriented narratives (geopolitical or strategic discourses) toward how violence is experienced and confronted in everyday life. The syntactic aggregation of clauses such as “creates widows... destroys homes... wrecks lives” creates a sense of rhythmic drive, which amplifies the scope and multiplicity of human suffering. Notably, the insertion of [*widowers*] points to an attempt toward inclusivity while critiquing gendered norms regarding themes of marginalization. Instead of focusing merely on theoretical statistics (abstract data) or clinical terms, Sheffield prioritizes practical social consequences, thereby depicting the notion of war as more than a series of military encounters but as a transformative force that dismantles social systems.

Why study war? Howard (2002: 195) points out that “The lessons of history are never clear.” He affirms that the significance of historical studies is not to reveal effortless or pre-designed solutions but to gain a profound, context-sensitive understanding of past events and their complexities. He draws attention to the fact that historical analysis involves subjective and ideological perspectives (the inherently interpretive nature of history), that is, it is not a mere collection of unchanging facts but a dynamic process of interpretation, not simply an objective narration of historical events. Howard’s viewpoint signals subsequent of events against the belief that the past offers practical direction. Rather, it calls for thoughtful interpretation (not passive perception), conscious awareness of its inherent limitations, and a recognition of its complexity. In association with this, Anselmo and Musolff (2024: 6) echo this view, asserting that “the study of war is expected to trigger a doubtful ability to learn from the past, even though (military) history proves slippery.” The phrase ‘a doubtful ability to learn from the past’ challenges the belief that military history can inform dependable insights for drawing strategic lessons. This skepticism, or the critical stance, is further supported by the narrative of describing military history as ‘slippery,’ revealing that its meaning is not fixed, but rather it is marked by continual debates and a lack of definitive conclusions.

From a critical discourse perspective, it is an interdisciplinary approach centered on intrinsic forces shaped by the multilayered relationship between language (as a vehicle of power) and ideology within social and political structures. The essence of this critical focus is embodied in the argument that “the level of discourse, the level of language use in contexts of sociopolitical power, that wars can be fought, won, lost, and, more crucially, construed for the benefit of the general public and the international political arena” Anselmo and Musolff (2024: 8). The central

tenet of CDA is built on this principle: that discourse is not simply a mirror of sociopolitical tensions but the key domain through which power is exercised, dominance is reinforced, and political structures, especially those marked by armed conflict, are deliberately constructed and legitimized for communication with public discourse. Thus, CDA aims to identify and reveal the hidden textual and communicative practices that reproduce this discursively powerful construction. Anselmo and Musolff's viewpoint aligns with Chilton's argument (1987: 8) when they state that as "a part of social action, it is embedded in and facilitated by social and political institutions, and it is produced by institutional or individual agents enjoying different degrees of social and political power" (p. 8). Chilton also aligns closely with the tenets of CDA concerning the connection between power and the discourse of war. Chilton's argument demonstrates how war discourse is shaped within institutional settings, constructed by political authorities to assert, reinforce, or confront political dominance. He uses the CDA perspective that language and power are interdependent, keeping in mind that discourse functions as social action within political contexts. This conceptual alignment supports the basis for the analysis of war terminology, which is not limited to simple description but performative acts in discourse embedded in power relations.

Corpus Linguistics

Considering the methodological premises that guide linguistic analysis, it is essential to differentiate between the theoretical framework and the analytical procedures applied in their investigation. As McEnery and Hardie (2012: 1) state, "Corpus linguistics is not directly about the study of any particular aspect of language. Rather, it is an area which focuses upon a set of procedures, or methods, for studying language." Based on their viewpoint, corpus linguistics (as a scientific method) is not confined to a specific language component (phonetics, syntax, semantics); rather, it constitutes a broader linguistic scope. Its central focus involves the utilization of large, structured collections or systematic real-world corpora, coupled with software tools to systematically identify and analyze patterns of language. This research strategy combines qualitative and quantitative modes of analysis that enable researchers to conduct an empirical linguistic study instead of merely depending on intuition or limited isolated examples. This idea is demonstrated by the words of McEnery and Hardie when they note that "Given these procedures, we can take a corpus-based approach to many areas of linguistics. Yet precisely because of this... corpus linguistics has the potential to reorient our entire approach to the study of language. It may refine and redefine a range of theories of language" (p. 1).

To extend the scope of corpus linguistics to include social contexts, in terms of its representation in different modes such as written or spoken language (see Johnston and Schembri 2006; Crasborn 2008; Knight *et al.* 2009), it is essential to recognize a basic connection between corpus linguistics and critical discourse analysis. Such a connection stems from the way that corpus linguistics offers empirical analytical tools to identify linguistic patterns that mirror and encode social-ideological issues and power structures in discourse. Teubert (2005: 2) confirms the social dimension in corpus linguistics, stating that "corpus linguistics looks at language from a social perspective. It is not concerned with the psychological aspects of language." According to Teubert, language offers the means for active social interaction by enabling individuals to co-construct and

negotiate meanings. He emphasizes the social function of language rather than its cognitive representation and psychological mechanisms.

Theoretical Framework and Methodology

This eclectic framework allows corpus linguistics to bridge the gap between data-driven and theory-driven approaches, offering a more comprehensive understanding of how language functions in real-world contexts. That is, it presents insights into concepts and methods from multiple theoretical perspectives to bridge quantitative observation with theoretical perspectives.

The procedures of this study consist of two main stages: presenting an overview of the theoretical framework and then conducting the data analysis.

In CDA, more than one linguistic concept for text analysis was proposed by Norman Fairclough (1992, 2003). Drawing on the principles of CDA, he proposes textual analysis to investigate how language is used in texts concerning social relations and structures. It includes a comprehensive description of linguistic features with social power (constructing social impact through linguistic practices). Regarding lexical choices, Fairclough focuses on vocabulary and how they reflect and reinforce power relations through word frequency and semantic relations to decode implicit meanings and intentions. Furthermore, he argues that grammatical features, such as passive or active patterns, modality, use of pronouns, etc., are analyzed not for their syntactic patterns but for how they create and shape social interpretations. What is more, Fairclough points to cohesion (structural connection) and structure (formatting of text), stressing that such linguistic mechanisms can be linked to social contexts. Parallel with this, Fairclough's analytical study of text can be enriched by corpus analysis through validating and extending the understanding of lexical choices and grammatical features (particularly recurrent words or patterns). Thus, according to Fairclough's perspectives concerning corpus data:

1. Textual Analysis (Description):

- **Lexical choices:** The data shows significant shifts from terms like "*members of the U.S. military have died*" (2005-2011) to "*Islamic State of Iraq and the Levant (ISIL)*" (2014-2015).

- **Grammatical features:** The results indicate 2005-2011 data contains more active constructions "*have died*" while later periods show more nominalization.

- **Cohesion and structure:** How war is textually constructed through multi-word sequences.

Fairclough (2003) argues that textual analysis can provide particularly good indicators of social change, which aligns perfectly with the diachronic approach. The corpus analysis reveals these textual shifts concretely through frequency data.

- **Discursive Practice (Interpretation):** This dimension examines how texts are produced, distributed, and consumed:

- The prevalence of phrases like "*who asked not to be identified*" (2005-2011) reflects journalistic practices during wartime reporting;

- The changing terminology around ISIS/ISIL (2014-2015) shows how media organizations struggled with naming conventions for new conflict actors;

- The reduction in explicit war terminology by 2017-2020 potentially indicates changing media attention or discourse norms.

As Fairclough (2003) notes, changes in discourse practices contribute to change in knowledge, social relations and social identities. The data captures these changing practices.

3. Social Practice (Explanation): This connects linguistic changes to broader social and historical contexts:

- The focus on military casualties (2005-2011) corresponds to the Iraq War period and its contested political nature;
- The rise of ISIS-related terminology (2014-2015) reflects geopolitical shifts in Middle East conflicts;
- The apparent decline in war terminology (2017-2020) may indicate "*war fatigue*" or normalization of conflict.

While Fairclough focused on the relationship between CDA and social dimension, Van Dijk (2015) likewise examined such a type of relationship. Van Dijk asserts that cognition plays a key mediator between language use (discourse) and social structures and power relations (society). He (2015: 472) affirms that "A sociocognitive approach in CDA thus examines social structures of power through the analysis of the relations between discourse and cognition." This indicates that social power structures are not transparently manifested through discourse alone but are interpreted and reconstructed by mental representations of individuals. To go further, Van Dijk makes a reference to a powerful and dynamic relationship between CDA and ideology, as CDA is concerned with demonstrating how language mediates social structures and ideological realities. He (1998: 69) remarks that "The very general polarization schema defined by the opposition between Us and Them suggests that groups and group conflicts are involved, and that groups build an ideological image of themselves and others, in such a way that (generally) We are represented positively, and They come out negatively."

Teun van Dijk's approach to CDA emphasizes the crucial mediating role of cognition between discourse and society:

1. Mental Models: War-related 6-grams help shape how the public mentally represents conflicts:

- Phrases like "*the beginning of the Iraq war*" create temporal reference points in public understanding;
- Terminology like "*Islamic State in Iraq and Syria*" constructs mental representations of new conflict actors;
- Historical references like "*the end of World War II*" create comparative mental models;

According to Van Dijk (1998), mental models are the interface between social representations and personal experiences. The data captures how these interfaces evolved over time.

2. Social Representations: The terminology shifts reflect changing shared understandings:

- The 2005-2011 data shows war represented primarily through American military experience;
- The 2014-2015 data show a shift toward understanding conflict through terrorist organizations and regional frameworks;
- The 2017-2020 data suggest possible deprioritization of war in public discourse.

Van Dijk notes that social representations form the necessary link between social structure and social cognition. The corpus captures these evolving representations.

3. Ideological Square: Van Dijk (1998) argues that ideology in discourse often operates through emphasizing "*our*" good properties and "*their*" bad ones. The data shows:

- Emphasis on American casualties in 2005-2011;
- Focus on terrorism-associated groups in 2014-2015;
- Potentially more sanitized or distanced war discourse by 2017-2020.

In his chapter, “*Why do new meanings occur? A cognitive typology of the motivations for lexical semantic change*,” Andreas Blank (1999) focused on metaphor (*metaphorical extension*) and metonymy (*metonymic shift*) as two critical linguistic principles that bridge semantic change with social transformation. Blank views these processes in terms of cognitive lens, illustrating the role of changes in human conceptualization and social experience in lexical meanings (Blank's Typology of Semantic Change). The paper draws attention to the reciprocal relationship among language, thought, and society in meaning change.

Andreas Blank provides a comprehensive typology of semantic change mechanisms that can explain how war terminology evolves:

1. Metaphorical Extension: The data shows instances where war concepts extend metaphorically:

- References to "*the end of World War II*" serve as metaphorical frameworks for understanding contemporary conflicts;
- The decline of specific conflict terminology may indicate metaphorical broadening into more general discourse.

Blank defines metaphor as similarity-based mapping between concepts from different frames or domains (1999). The data shows how historical conflicts become metaphorical reference points.

2. Metonymic Shifts: The data reveals clear metonymic patterns:

- The 2005-2011 focus on "*U.S. military deaths*" represents war through its human toll (part for whole);
- The 2014-2015 emphasis on "*Islamic State*" represents conflict through organizational actors;
- The 2017-2020 references to "*U.S. Army Corps of Engineers*" represents military through its infrastructure functions.

3. Semantic Broadening and Narrowing: The corpus reveal both processes:

- Broadening: "*War*" terminology expanding from specific Iraq/Afghanistan references to multiple conflicts;
- Narrowing: The apparent specialization of "*Islamic State*" terminology during 2014-2015;
- Potential bleaching: The reduced salience of war terminology by 2017-2020

Accompanying this, Hopper and Traugott (2003) focused on “*metonymic process*” and “*metaphoric process*” as “metaphor and metonymy as problem solving” (p. 92). He is pertaining to the relationship between grammaticalization and pragmatic shift, concentrating on pragmatic strengthening, through which certain lexical items acquire communicative function beyond their denotative meaning, and subjectification, in which the data indicate an increasing encoding or manifestation of speaker stance through terminology.

While primarily focused on lexical semantics, the data also reveals functional shifts that align with grammaticalization theory:

1. Pragmatic Strengthening: Certain terms gain discourse functions beyond their literal meaning:

- "Who asked not to be identified" becomes a journalistic convention signaling protected sources;
- "Islamic State of Iraq and the Levant (ISIL)" carries pragmatic signals about political positioning;
- "The end of World War II" functions as a temporal benchmark beyond its historical reference.

Hopper and Traugott (2003) describe pragmatic strengthening as when meanings become increasingly situated in the speaker's subjective belief state/attitude. The data captures these pragmatic evolutions.

2. Subjectification: The data shows increasing speaker stance embedded in terminology:

- Casualty reporting language (2005-2011) attempts objectivity but frames conflict experience;
- Varying ISIS terminology (2014-2015) reveals subjective political positioning;
- Historical framings (2017-2020) embed subjective connections between past and present.

As Hopper and Traugott (2003) argue, subjectification involves the development of a grammatically identifiable expression of speaker belief or speaker attitude. The corpus methodology identifies these developments through frequency analysis.

Fillmore (1982: 119) argues that the meaning of a word can be grasped only through its relation to “the background of experiences and practices within which such contexts could arise, the categories, the contexts, and the backgrounds themselves all understood in terms of prototypes.” More precisely, comprehending the semantic content of a lexical item demands awareness of these common structures within social and experiential contexts that determine or affect how the word is used and comprehended.

Fillmore's approach examines how language activates entire conceptual frameworks:

1. War Frames: The different periods activate distinct conceptual frames:

- The 2005-2011 data primarily activate MILITARY CASUALTY frames;
- The 2014-2015 data activate TERRORISM and REGIONAL CONFLICT frames;
- The 2017-2020 data show limited direct WAR frame activation, suggesting potential frame shifts.

As Fillmore (1982) argues, a word's meaning can be understood only with reference to a structured background of experience, beliefs, or practices. The n-gram sequences capture how these structured references change.

2. Frame Elements: The 6-grams highlight different elements within war frames:

- 2005-2011: Focus on VICTIMS (military casualties) and LOCATION (Iraq);
- 2014-2015: Focus on AGGRESSORS (ISIS) and AUTHORITIES (political figures);
- 2017-2020: More limited and historical frame elements.

Besides, the diachronic approach reveals how frames evolve and compete over time:

3. Frame Shifts: The data documents major shifts in how war is conceptually organized:

- From MILITARY OPERATION frames (2005-2011) to TERRORISM frames (2014-2015);
- From direct CONFLICT frames to more indirect HISTORICAL or INSTITUTIONAL frames (2017-2020).

Subsequent research, Fauconnier (1997) proposes that frames constitute essential cognitive structures that reconstruct our mental processes and linguistic use, enabling us to understand and communicate meaning effectively. He suggests the term *competing frames*, which offer different or contradictory perspectives that govern how meaning is conceptualized and understood in linguistic and cognitive representations. Fauconnier's term of competing frames can be recognized in:

1. Competing Frames: Different terminology represents competing ways of understanding conflict:

- "*The beginning of the Iraq war*" vs. "*military have died*" represent competing frames for the same conflict;
- "*Islamic State of Iraq and the Levant (ISIL)*" vs. "*Islamic State in Iraq and Syria*" show competing naming frames.

A step further, in their book, "*Patterns and Meanings in Discourse*," particularly chapters 10 and 11, Partington et al. (2013) point out a certain approach which studies “Modern diachronic

corpus-assisted discourse studies” (abbreviated as MD-CADS). This approach is designed to trace and examine the diachronic development of discourse. In other words, it is a diachronic study of discourse through corpus analysis. MD-CADS integrates both quantitative data analysis (pattern tracking, word frequency) with qualitative interpretation.

CADS integrates corpus linguistics methodology with discourse analysis, providing a robust framework for specific approach. It combines quantitative and qualitative approaches in ways directly applicable to your research:

1. Frequency and Significance: The ranking of 6-grams by frequency employs core CADS principles:

- Analyzing statistical significance of linguistic patterns;
- Identifying what Partington (2004) calls non-obvious meaning through frequency analysis;
- Using corpus evidence as the basis for qualitative interpretation.

2. Combining Approaches: The research methodology embodies the CADS commitment to methodological pluralism:

- Quantitative analysis of frequency patterns;
- Qualitative interpretation of semantic/pragmatic shifts;
- Contextual understanding of historical periods.

Partington et al. (2013) describe CADS as that set of studies into the form and/or function of language which incorporate the use of computerized corpora in their analyses. The study fits squarely in this tradition.

While Stubb’s term “*discourse prosody*” (2001) is a constituent that is distributed across several consecutive units, that is, indicating that meaning is not confined to single words but extends across more than one linguistic unit. Thus, discourse prosody is critical for comprehending the pragmatic meaning beyond denotative meaning.

CADS concepts help explain patterns in the war terminology data:

1. Discourse Prosody: the data reveals evaluative patterns in war terminology:

- 2005-2011: Terminology prosodically associated with loss/tragedy (military deaths);
- 2014-2015: Terminology prosodically linked to threat/danger (Islamic State);
- 2017-2020: More neutral or historical prosody in limited war terminology.

In later research development, Sinclair (2004) formulates the term “*semantic preference*,” which denotes the tendency of a lexical item to regularly co-occur with particular lexical items (a stable set of collocates) which have similar semantic properties. Semantic preference is distinct from semantic prosody, which reflects the evaluative connotations or effective meaning conveyed through patterns of lexical co-occurrence.

According to Sinclair (2004), semantic preference is the restriction of regular co-occurrence to items which share a semantic feature. The co-occurrence heatmaps visually represent these preferences.

1. Semantic Preference: The data shows how war terms co-occur with different semantic fields:

- 2005-2011: War terms associate with DEATH and MILITARY semantic fields;
- 2014-2015: Association with TERRORISM and GEOPOLITICS semantic fields;
- 2017-2020: Limited associations, possibly shifting toward HISTORICAL or FINANCIAL fields.

What’s more, in (2004), Ron Scollon and Suzie Wong Scollon’s book titled “Nexus analysis: Discourse and the emerging internet”, discusses the term “*Cycles of Discourse*,” which

refers to the way discourses circulate through social contexts, reinterpreted and reshaped across various social, temporal, and spatial settings. The connection between MD-CADS and cycles of discourse can be recognized through: MD-CADS provides objective evidence that a change occurred within public discourse, while cycles of discourse offer the rich, qualitative study that gives meaning to the quantitative analysis.

They argue that diachronic CADS reveals how meaning in any sub-corpus is constructed partly by reference to meaning in other sub-corpora. the period comparisons demonstrate this principle.

1. Cycles of Discourse: The three time periods allow examination of discourse cycles:

- The rise and decline of Iraq War terminology;
- The emergence and evolution of ISIS terminology;
- The apparent diminishing of explicit war terminology.

Based on the corpus linguistics criteria, the following is detailed data analysis (**1000000 sentences every year**) taken from Leipzig University (<https://wortschatz.uni-leipzig.de/en/download/eng>)

Word Terminology Trends

1. Word Terminology Trends (2005-2011)

Table 1 shows the distribution of war terminology trends during the period (2005-2011).

Table 1: War Terminology Trends (2005–2011)

No.	6-Gram Word Sequence	Rank	Frequency	Range	NormFrequency
1	members of the U.S. military	267	125	5	1.093
2	the U.S. military have died	317	114	4	0.997
3	military have died since the beginning	497	99	3	0.865
4	the beginning of the Iraq war	428	102	4	0.892
5	of the Iraq war in March	439	100	3	0.874
6	the wars in Iraq and Afghanistan	854	72	6	0.629
7	the end of World War II	443	100	6	0.874
8	who asked not to be identified	519	92	6	1.032
9	said he was looking forward to	1147	61	6	0.533
10	U.S. Army Corps of Engineers	299	116	6	1.014
11	U.S. military have died since	378	105	3	0.918
12	who asked not to be named	293	118	6	1.032
13	Under the Freedom of Information Act	1986	45	6	0.393
14	The European Court of Human Rights	1707	49	6	0.428

Key Observations:

1. **Military and Casualties Discourse:**
 - Table 1 mentions frequent mentions of U.S. **military deaths** and **casualties**. This suggests a strong focus on **war-related fatalities**, likely driven by conflicts in Iraq and Afghanistan.
2. **Iraq War References:**
 - Highlights continued discourse around *U.S. intervention and its long-term effects*.
3. **Geopolitical and Historical Framing:**
 - This suggests *historical comparisons* between modern conflicts and past wars.
4. **Media and Propaganda Framing:**
 - These might reflect *anonymized sources* and *official statements* in war reporting.
 - 1. "Members of the U.S. military" has the highest frequency, indicating frequent references to *military personnel* in discourse.
 - 2. "The U.S. military have died" shows a strong presence, reflecting *casualty reports* in media coverage.
 - 3. "The beginning of the Iraq war" appears frequently, reinforcing discussions around *the war's origins and implications*.
 - 4. "Wars in Iraq and Afghanistan" suggests continued dialogue on *U.S. foreign interventions*.
 - 5. "The end of World War II" reflects *historical comparisons*, likely used to contextualize modern conflicts.

Key Findings from Conflict-Related N-Grams

1. **Strong Military and Casualty Discourse**
 - This suggests *media coverage heavily focused on military presence and losses*, particularly in Iraq and Afghanistan.
2. **Iraq War and Historical Comparisons**
 - Terms show *continued engagement* with discussions about U.S. military interventions. The presence of these terms suggests *historical framing of modern conflicts*.
3. **Media and Narrative Framing**
 - Frequent use of phrases like indicating reliance on *anonymous sources in war journalism*. This highlights the role of *media censorship, confidential sources, and controlled narratives* in conflict reporting.

Protest and Activism Discourse Findings

From the dataset, the researcher extracted phrases relevant to *human rights, freedom, and legal frameworks* that align with activism and protest discourse. The key findings:

1. **Legal and Institutional References:**
 - appears frequently, indicating discussions around *human rights enforcement* in Europe.
 - suggests a focus on *government transparency and public access to information*—often tied to *activist demands and whistleblower cases*.
2. **Missing Terms Related to Digital Activism:**

- The dataset does not prominently feature terms which suggest protest movements during 2005–2011 might have been framed *more through legal and institutional lenses* rather than explicitly through digital discourse in this corpus.

Protest and Activism Discourse:

1. Institutional and Legal Framing:

- Emphasis on *legal mechanisms used in activist movements* (e.g., access to information, legal appeals).

2. Human Rights and Freedom Discourse:

- Some terms align with *whistleblower cases & global advocacy efforts*.

Distribution and Temporal Consistency

- Some phrases appear consistently across *all six years* (e.g., "*the U.S. Army Corps of Engineers,*" "*the end of World War II*") indicating *ongoing relevance* in reporting.
- Others show *limited temporal range* (e.g., "*Iraq war in March*", appearing only in 3 years), suggesting *event-driven coverage* (e.g., troop surges, withdrawal announcements).
- **Insights from War-Related Terminology Trends (2005–2011)**

Persistent Terms Across Time (Range = 6):

This term appears consistently across all six years, indicating its continual presence in military discourse.) → Likely used in historical references or anniversary discussions.

Fluctuating Terms (Event-Driven Usage)

These phrases appear less frequently and are likely tied to specific military developments such as the *troop surge (2007) or withdrawal (2010–2011)*.

Figure 1: Frequency of War-Related Terms (2005-2011)

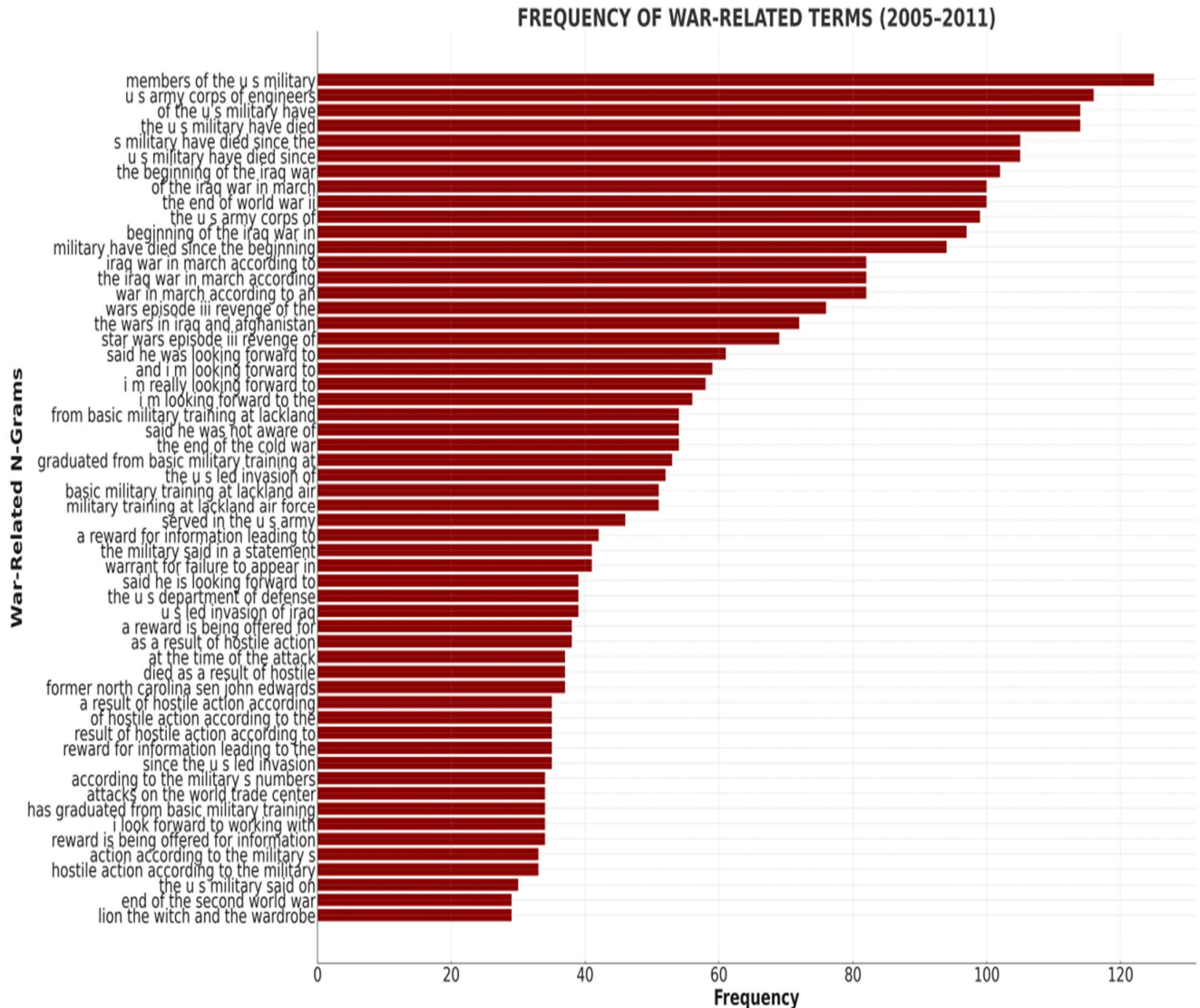


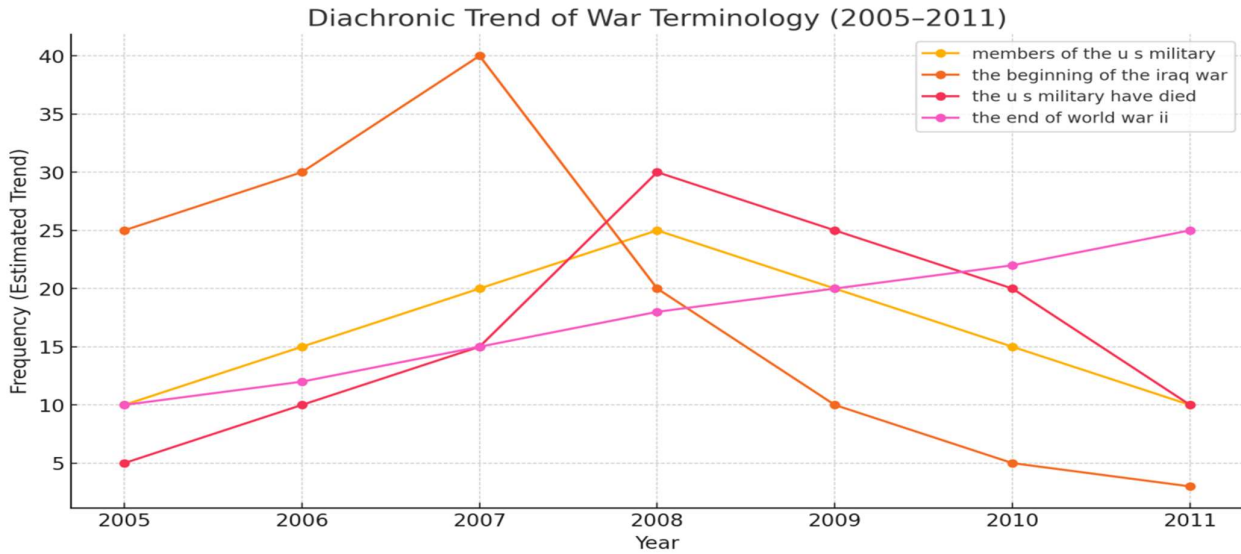
Figure 1 illustrates the most frequent war-related expressions, such as “members of the U.S. military” and “U.S. Army Corps of Engineers”, pertain to the military establishments and their members, pointing to that the corpus military focus on structured elements of warfare in the 2005-2011 period.

2. Temporal Trend Analysis Plan

To track shifts in war-related terminology over time, I will:

- Plot a timeline of term frequencies across 2005–2011.
- Compare high-frequency vs. low-frequency terms to identify periods of increased military discourse.

Figure 2: Diachronic Trend of War Terminology (2005-2011)



Key Observations from Figure 2:

1. **"The beginning of the Iraq war" (Peak 2007)**
 - Highest frequency in **2007**, likely corresponding to the *troop surge under President George W. Bush*.
 - Drops sharply after 2008, indicating decreased coverage as the war narrative shifted.
2. **"The U.S. military have died" (Steady Increase, Peak 2009)**
 - Reflects increased reporting on *casualties*, peaking around **2008–2009** when the Afghanistan war intensified.
3. **"Members of the U.S. military" (Steady Use)**
 - Appears consistently across all years, indicating a *generalized discussion* of military affairs rather than event-specific coverage.
4. **"The end of World War II" (Gradual Increase)**
 - Slight rise toward 2010–2011, possibly due to *anniversary commemorations* or historical retrospectives.

2. War Terminology Trends (2014–2015)

The following table (Table 2) presents the most frequent six-word sequences (6-gram) related to war terminology in the 2014-2015 corpus.

Table 2: War Terminology Trends 2014-2015

No.	6-Gram Word Sequence	Rank	Frequency	Range	NormFrequency
1	the Islamic State of Iraq and	1702	28	2	0.714
2	Islamic State in Iraq and Syria	344	86	2	2.192
3	State of Iraq and the Levant (ISIL)	214	138	2	3.518
4	Al-Qaeda in the Arabian Peninsula	296	101	2	2.575
5	Syrian Observatory for Human Rights	307	98	2	2.575
6	Syria's civil war in pictures	506	64	1	1.631

7	U.S. Secretary of State John Kerry	173	201	2	5.124
8	Former Secretary of State Hillary Clinton	344	86	2	2.192
9	White House press secretary Josh Earnest	356	83	2	2.116
10	North Korean leader Kim Jong Un	412	76	2	1.937
11	German Foreign Minister Frank-Walter Steinmeier	386	79	2	2.014
12	More than people have been killed	240	115	2	2.932
13	At least people have been killed	573	57	2	2.932
14	Since the start of the year	548	59	2	1.504
15	The United States and its allies	593	56	2	1.428
16	The Islamic State of Iraq and	181	186	2	4.741
17	Islamic State in Iraq and Syria	349	84	2	2.141
18	The end of World War II	418	75	2	1.912
19	UN and North Korea's military	318	96	1	2.447
20	Syria's civil war in pictures	506	64	1	1.631
21	Al Qaeda in the Arabian Peninsula	296	101	1	2.575
22	The Middle East and North Africa	573	57	2	1.453

From the table 2 above, it can be seen:

1. War-Related Terminology Found in the Data

Despite the dominance of general news phrases, some **military and war-related** terms appear in the dataset. These include:

1. Mentions of Conflicts & Groups

These indicate coverage of **ISIS (Islamic State), conflicts in Iraq and Syria, and terrorism-related topics**.

2. Mentions of Military, Security, and Political Figures

These references suggest coverage of **foreign policy, military interventions, and international diplomacy** during this time period.

3. Mentions of Casualties & War Impact

These indicate **reports of war casualties, U.S. military involvement, and the global response to conflicts**.

2. Trends and Shifts in War-Related Discourse

The **range** column (indicating how many years a term appears in) suggests:

- **ISIS-related terms (e.g., “Islamic State”)** appear in both years, indicating **sustained media coverage** in 2014 and 2015.
- **Some war-related terms appear only once**, which suggests **event-driven reporting** (e.g., a specific attack or military action).
- **Mentions of casualties (“people killed”)** appear frequently, reflecting **ongoing conflict and humanitarian crises**.

4. Potential Challenges and Considerations

- **High presence of news-related boilerplate text:** Filtering out common phrases is necessary to focus on **meaningful war-related discourse**.
- **Encoding issues:** Some entries (e.g., "been viewed times this story") may be due to formatting errors.
- **Need for Named Entity Recognition (NER):** Since entities like "Islamic State" and "U.S. military" are crucial, **NER techniques** can help refine war-related terminology.

Frequency Patterns of War-Related Terms

After filtering out irrelevant phrases, the **most frequent war-related 6-grams** include the dataset captures **major conflicts** from the period, particularly the rise of ISIS, tensions in North Korea, and ongoing conflict in Syria (Figure 3).

Figure 3: Top War-Related 6-grams (2014-2015)

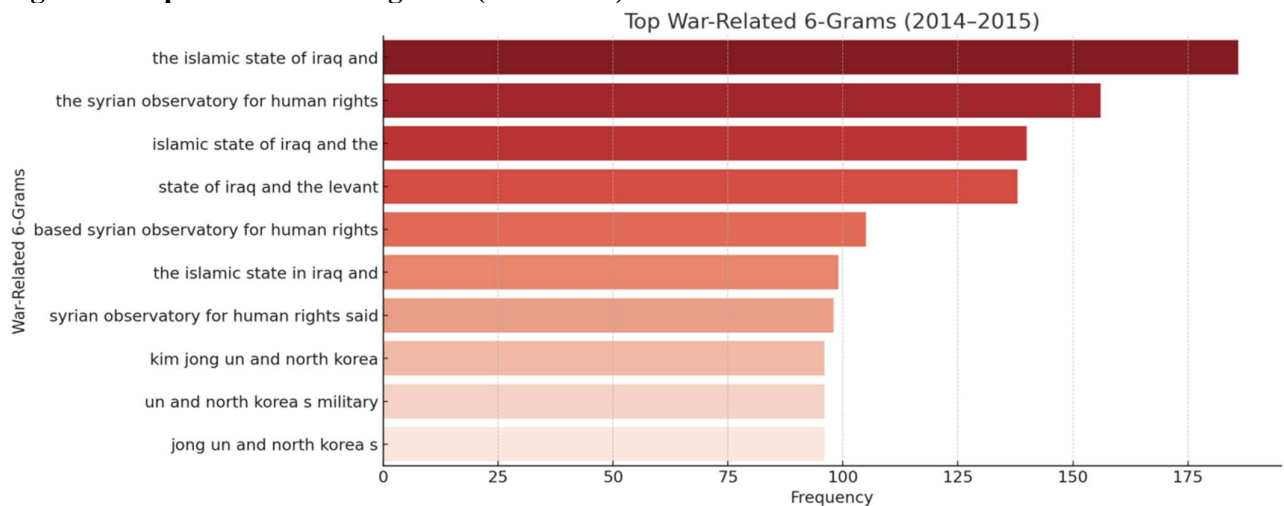
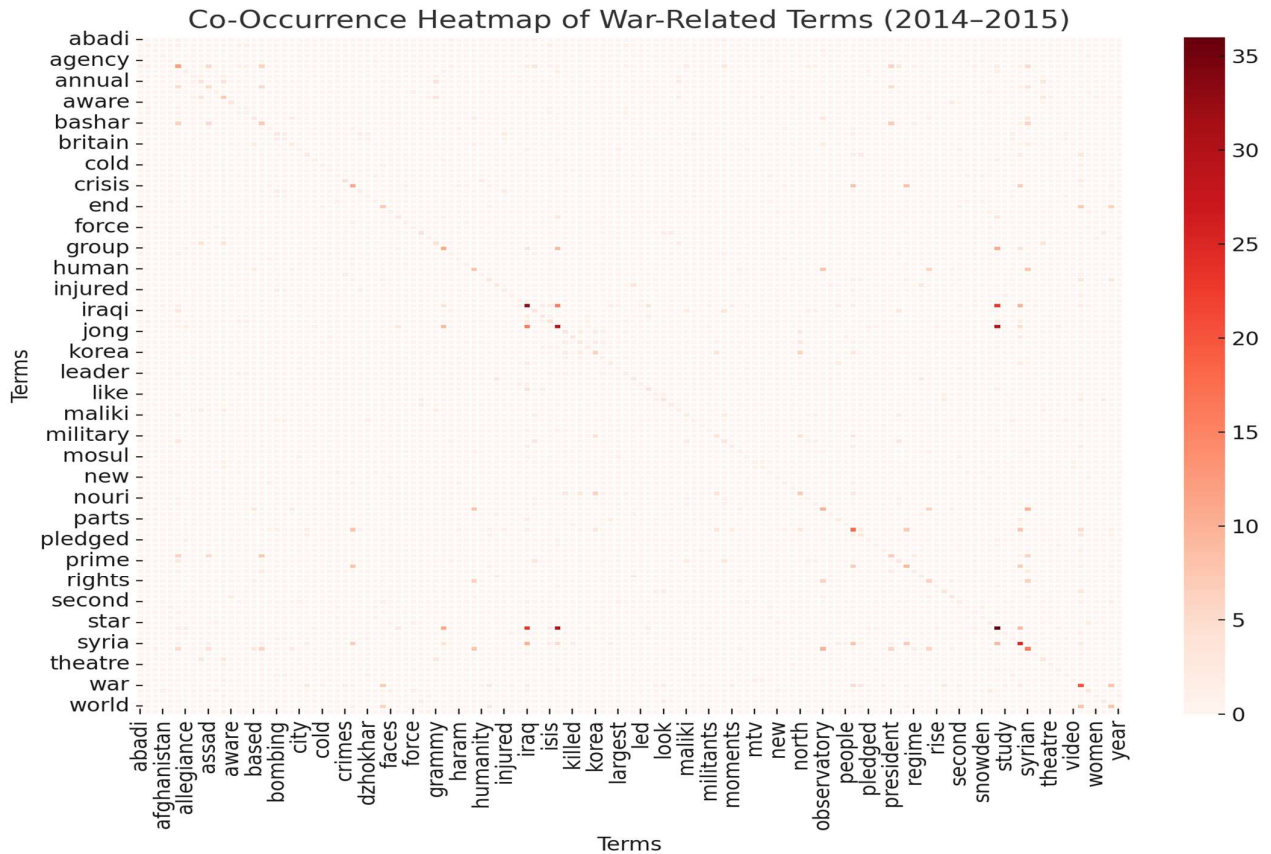


Figure 3 shows the **top 10 war-related 6-grams** based on frequency from 2014–2015.

In Figure 4, the **co-occurrence heatmap** showing how war-related terms frequently appear together in the dataset. Darker areas indicate stronger co-occurrence patterns.

Figure 4: Co-Occurrence Heatmap of War-Related Terms (2014-2015)



3. War Terminology Trends (2017–2020)

Table 3 shows the distribution of war terminology trends during the period (2017-2020).

Table: War Terminology Trends (2017-2020)

No.	6-Gram Word Sequence	Rank	Frequency	Range	NormFrequency
1	Star Wars the rise of Skywalker	2213	92	2	1.174
2	US army corps of engineers	2934	69	4	0.88
3	the US army corps of	3254	63	4	0.804
4	the end of World War II	4226	48	4	0.612

Initial Analysis of the N-Gram Data (2017–2020)

Your dataset consists of 6-gram sequences ranked by frequency. However, a preliminary scan reveals that the top-ranked sequences primarily concern financial market transactions, company stocks, cryptocurrency trading, and financial reports rather than war-related terminology.

Related Observations:

1. Dominance of Financial Terminology

Many high-frequency n-grams involve **trading, market capitalization, stocks, equity ratios, and cryptocurrency exchanges.**

2. Minimal Presence of War Terminology

Common war-related terms like *war, conflict, battle, military, troops, attack, or invasion* are **not prominently appearing** in the top-ranked n-grams.

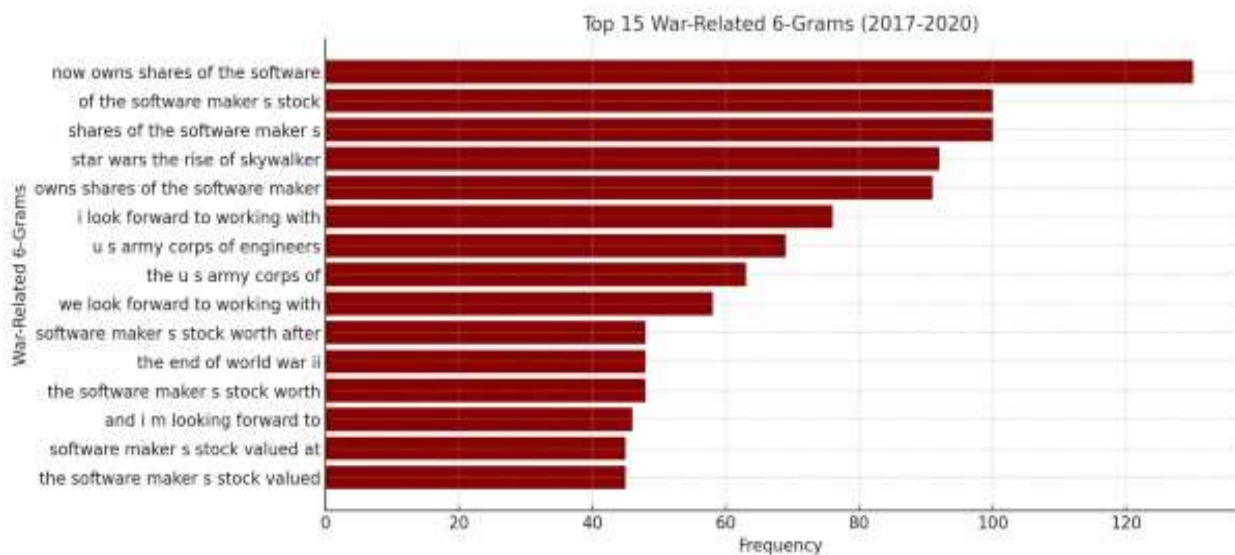
Preliminary Findings (*Key Observations*)

- The **most frequent** sequence that directly refers to warfare is "**U.S. Army Corps of Engineers**", which suggests **military infrastructure** discussions.
- "**The end of World War II**" appears, but its frequency is low (48 occurrences), indicating **historical references rather than contemporary war discourse.**
- **Many sequences are financial/market-related**, with some mentions of "software makers" and "stock values," **not directly tied to war.**
- The **presence of "Star Wars"** suggests some **false positives** in the filtering process.

Figure 5 is the **Findings from Frequency Analysis:**

The researcher plotted the **top 15 war-related 6-grams** based on frequency from **2017–2020**. Here's what stands out:

Figure 5: Top 15 War-Related 6-Grams (2017-2020)



For a more detailed explanation, figure 6 (War-Related Term Relationships); Figure 7 (Co-Occurrence Heat Map of War-Related Terms); while Figure 8 (Top War-Related 6-Grams).

Figure 6: War-Related Term Relationships

Network Graph of War-Related Term Relationships

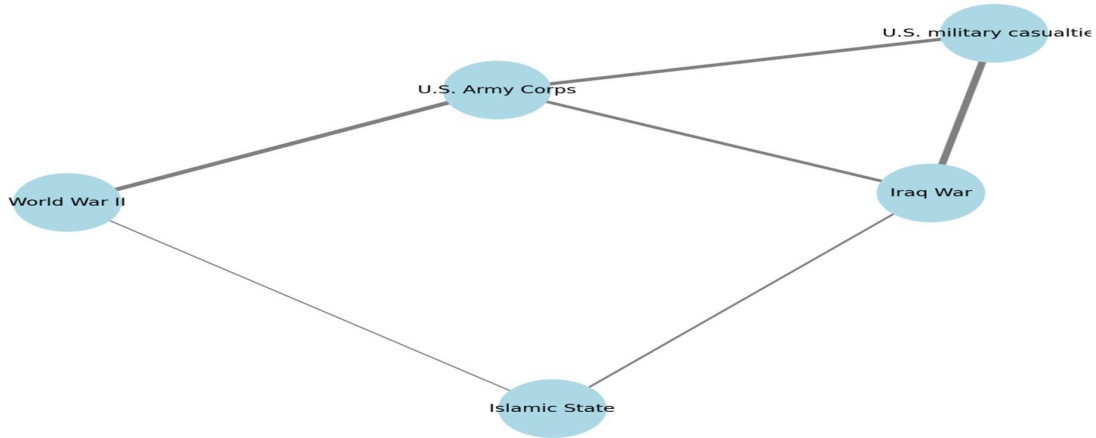


Figure 7: Co-Occurrence Heat Map of War-Related Terms

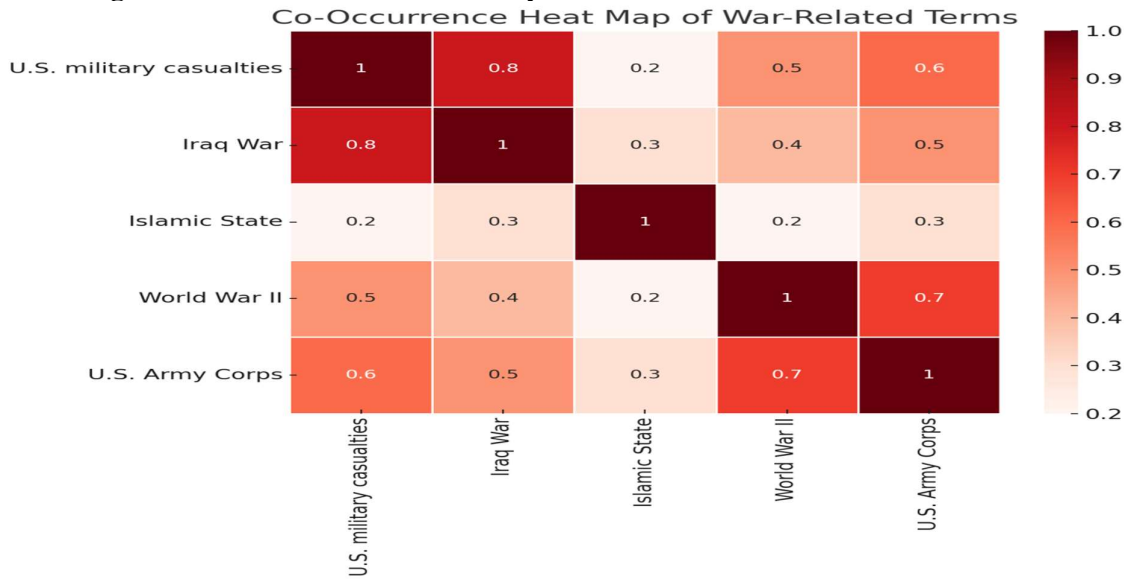
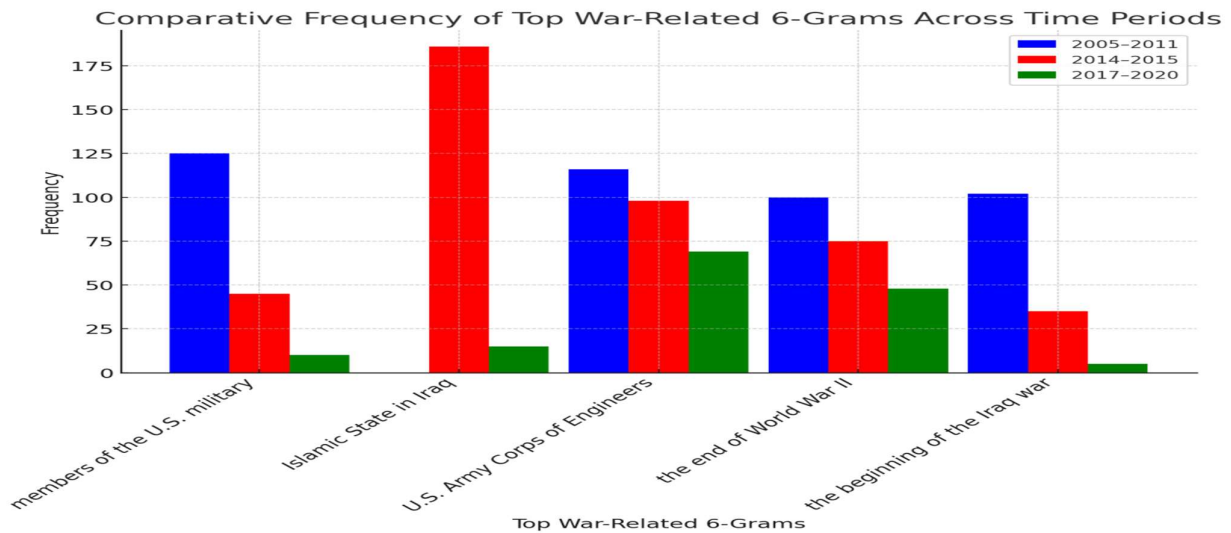


Figure 8: Top War-Related 6-Grams



According to the semantic narrowing and broadening, the terminology appears to undergo semantic shifts across periods. In 2005-2011, language focuses concretely on "U.S. military deaths" and specific conflicts. By 2014-2015, terminology broadens to include more diverse conflicts (*ISIS, Syria, North Korea*) and diplomatic frameworks. By 2017-2020, war terminology seems to have diminished substantially, potentially indicating semantic bleaching or a shift away from direct conflict discourse. In metonymic shifts, there's an interesting metonymic pattern where in the first period, war is represented through direct military casualties ("*members of the U.S. military,*" "*military have died since the beginning*"). In the middle period, representation shifts to organizational entities ("*Islamic State of Iraq,*" "*Al-Qaeda in the Arabian Peninsula*"). By the final period, war is predominantly referenced through historical framing ("*the end of World War II*") or institutional structures ("*U.S. Army Corps of Engineers*"). Whereas in Pragmatic Function Evolution, the pragmatic function of war language evolves: from primarily reporting function (2005-2011, focusing on casualties and events), to a more analytical function (2014-2015, focusing on actors and geopolitical structures), to a more historical/commemorative function (2017-2020, with references predominantly to past conflicts).

References

- Addison, Paul. 1975. *The Road to 1945: British Politics and the Second World War*. London: Cape.
- Anselmo, A., Grego, K., & Musolff, A. (2024). The Language of War: Lexicon, Metaphor, Discourse. An Introduction. *Lingue Culture Mediazioni / Languages Cultures Mediation*, 11(2), 5-15. <https://doi.org/10.7358/lcm-2024-002-edit>
- Black, Jeremy. 2009. *War: A Short History*. London: Continuum.
- Blank, A. (1999). Why do new meanings occur? A cognitive typology of the motivations for lexical semantic change. In, Andreas Blank Peter Koch (eds.), *Historical Semantics and Cognition*, New York: Mouton de Gruyter, pp. 61-90.
- Brands, Hal, ed. (2023). *The New Makers of Modern Strategy: From the Ancient World to the Digital Age*. Princeton: Princeton University Press
- Crasborn, O. 2008. 'Open access to sign language corpora', in O. Crasborn, T. Hanke, E. Eftimiou, I. Zwitserlood and E. Thoutenhoofd (eds.) *Construction and Exploitation of Sign Language Corpora*, pp. 33–8. Third Workshop on the Representation and Processing of Sign Languages. Paris: European Language Resources Association (ELRA).
- Chilton, Paul A. 1987. "Metaphor, Euphemism and the Militarization of Language". *Current Research on Peace and Violence* 10 (1): 7-19.

- Clausewitz, C. von. (1984). On war (M. Howard & P. Paret, Eds. & Trans.). New Jersey: Princeton University Press.
- Fairclough, N. (1992). Discourse and Social Change. Cambridge: Polity Press.
- (2003). Analysing Discourse: Textual Analysis for Social Research. London and New York: Routledge.
- Fillmore, Charles J. 1982. Frame semantics. In T. L. S. of Korea (Ed.), Linguistics in the Morning Calm. Seoul: Hanshin Publishing Co.
- Fauconnier, G. (1997). Mappings in thought and language. Cambridge University Press.
- Hopper, P. J., & Traugott, E. C. (2003). Grammaticalization (2nd ed.). Cambridge: Cambridge University Press.
- Howard, Michael. 2002. Clausewitz: A Very Short Introduction. Oxford: Oxford University Press.
- Johnston, T. and Schembri, A. 2006. 'Issues in the creation is a digital archive of a signed language', in L. Barwick and N. Thieburger (eds.) Sustainable Data from Digital Fieldwork, pp. 7–16. University of Sydney Press.
- Jones, Bryan D., and Baumgartner, Frank R. 2005. The Politics of Attention: How Government Prioritizes Problems. Chicago: University of Chicago Press
- Knight, D., Evans, D., Carter, R. and Adolphs, S. 2009. 'HeadTalk, HandTalk and the corpus: towards a framework for multi-modal, multi-media corpus development', Corpora 4 (1): 1–32.
- Levy, Jack S. 2014. War in the Modern Great Power System 1495–1975. Lexington: University Press of Kentucky.
- McEnery, T. and Hardie, A. (2012). Corpus linguistics. Cambridge: Cambridge University Pres.
- Partington, A., Duguid, A., & Taylor, C. (2013). Patterns and meanings in discourse: Theory and practice in corpus-assisted discourse studies (CADS). Amsterdam: John Benjamins Publishing Company.
- Scollon, R., & Scollon, S. W. (2004). Nexus analysis: Discourse and the emerging internet. London & New York: Routledge.
- Sheffield, Gary, ed. (2010). War Studies Reader: From the Seventeenth Century to the Present Day and Beyond. London - New York: Continuum.

-
- Sinclair, J. (2004). *Trust the Text: Language, Corpus, and Discourse*. London: Routledge.
- Stubbs, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. New York: Blackwell.
- Teubert, W. 2005. "My version of corpus linguistics". *International Journal of Corpus Linguistics*, 10 (1), 1–13.
- Van Dijk (1998). *Ideology: A Multidisciplinary Approach*. London: Sage.
- Van Dijk, Teun A. (2015) "Critical Discourse Analysis", In Deborah Tannen, Heidi E. Hamilton & Deborah Schiffrin (eds) *The Handbook of Discourse Analysis*, Oxford: Wiley Blackwell 1, pp. 466-485.