



# Linnéuniversitetet

Kalmar Växjö

## Papa Revisited

A Corpus-Stylistic Perspective on the Style and Gender Representation of Ernest Hemingway's Fiction



Examiner: Per Sivefors

Authors:

Johan Nilsson

[jn222nn@student.lnu.se](mailto:jn222nn@student.lnu.se)

Daniel Sundberg

[ds222fd@student.lnu.se](mailto:ds222fd@student.lnu.se)

Supervisors:

Johan Höglund

Jukka Tyrkkö

HT-17

50EN01E

Abstract:

This essay revisits some of the more well-cited works of close to a century of scholarly and biographical efforts on the author Ernest Hemingway. It aims to re-evaluate and test the general assumptions, descriptions and specifications of his textual style and depiction of women through modern corpus stylistic methods. Through parallels between contextual material and periods of publication this project will explore the degree to which the common assumptions and descriptions of Hemingway's fiction hold true, and to which degree they can legitimately be treated as general descriptors of a literary style in development throughout a career of publication spanning a large part of the 1900's, both in terms of generalizations and definitions of changes taking place at specific times during the author's career. This essay will also define unresolved conflicts in the long history of Hemingway criticism and contribute towards finding an answer for the question of whether the descriptions could be considered generally correct, or defining the period of a description's relevance in regards to the author's published material. In the end, this essay intends to provide a further understanding of Hemingway's style, providing basis for new and more specific academic work on his authorship in the future.

Content:

Introduction.....	1
Stylistics as Method .....	5
Feminist Literary Criticism.....	12
The Specialized Corpus.....	15
Source Selection.....	18
Modes of Enquiring the Hemingway Corpus.....	20
Hemingway's Background and Influences.....	25
Hemingway's Style in Data.....	28
Gender Participation in Hemingway's Writing.....	40
Gender Representation in Hemingway's Writing.....	47
Discussion and Conclusion.....	56
Works Cited.....	62
Appendices:	
A: Hemingway Corpus Content.....	65

## Introduction

Ernest Hemingway has been the focus of numerous literary studies belonging to various fields, both today and during his own time. A vast amount of scholarly material has been published scrutinizing the various constituents of Hemingway's literary style. It is, of course, in the nature of a descriptive text such as a biography or a scholarly introduction to generalize main features of style, but sometimes these descriptions clash. As a result of this, descriptions of the Hemingway's style remain in conflict with each other due to the issues with providing a quantitative answer to whichever question lies behind the differing opinions. Of course, disagreements can be due to any number of other factors or differing interpretations, but the quantitative data here is meant to further the discourse rather than provide any irrefutable answers.

This essay revisits the current seminal works of scholarly and biographical efforts on Ernest Hemingway and aims to re-evaluate and test the general assumptions, descriptions and specifications of his textual style through modern corpus stylistic methods using computational analysis of linguistic features. This is done in an effort to examine Hemingway's language, descriptions of gendered participation and the representation of women in Hemingway's work with a base in the critical material. As a further aim, this essay will address unresolved conflicts in the long history of Hemingway criticism and contribute towards finding an answer for the question of whether or not the general descriptions could really be considered valid for the entirety of Hemingway's work.

Gathering these different perspectives, opinions and descriptions of Hemingway's style is central to this essay's aim. The selection must be authoritative, wide and encompass both the author's contemporaries, seminal works from the decades since his passing as well as a selection of influential recent research and opinions on the topic. It is also of significant importance to keep the focus narrow and to only engage with descriptions of Hemingway's

style that are transferable to queries that can be processed by the tools at our disposal. This dismisses general descriptions of content that are not visible through the language itself, as this essay deals only with language features. This is due to the need for quantifiable data required by our method, and the fact that interpretation of the works is not the purpose of this study. This essay does not dismiss the fact that interpretational or intertextual connections are both important parts of literary style, but we have chosen not to focus on these aspects due to the limited time available for this project.

A corpus-based method is based around a database of texts and a tool set used to formulate queries that can provide interesting data on language structure and word formations. This is done by gathering and sorting texts of the relevant type and then using different types of software to investigate different traits found within the texts. The main use of a corpus is to look at patterns in language. For instance, seeing which words collocate, or are found in connection to, a specific word can provide insight into the common connotations of that specific word as other words found within the same structure at a high frequency can tell us how the word is commonly used or what feelings the word is commonly written in relation to.

A corpus is one of the main tools of modern linguistics when researching traits specific to a language in terms of standard usage and dialectal usage. Corpus stylistics uses the data available from a corpus to further elaborate on a text, or to answer questions about the language style. The approach can also be used to discover functional words or structures that appear often enough for a pattern to appear. The selection of texts is normally made in accordance with criteria on language style such as type, dialect or time of production. Software is then used in order for the researcher to gather quantifiable data on different aspects of the language in the database.

The patterns can, similarly to how they reveal new information about language use, reveal a hidden discourse or theme in a novel, built by language parts too small or

unremarkable to be noticed through close reading, thus furthering the understanding of the text. These discourses, themes and other new data can then be used in combination with an additional interpretational theory to further develop criticism on the text.

This study utilizes two different concordance tools that provide an adequately wide range of options in regards to possible queries and questions as the queries the tools are capable of processing fit well with the needs of this project. The different types of queries are specified in the section on modes of enquiry. The software used are Laurence Anthony's freeware concordance tool AntConc created at Waseda University in Japan and Mike Scott's tool suit WordSmith v.4.0, developed at Liverpool University, England. Both tool kits will be discussed further in the section on method and theory. Other tools can be used for minor, specific queries, and will be specified in the relevant section. The primary source of data will be the Hemingway corpus created for this project. The Hemingway corpus consists of 69 novels and short stories with a word count of 567,596. The corpus is further specified in appendix A. The corpus contains copyrighted materials, and therefore cannot be made available outside of the findings presented in this essay. Some of the methods used in this study require the use of secondary corpora, or reference corpora, in order to function properly. The reference corpora will be specified and properly referenced when used.

In broad strokes, this study will engage with current works of Hemingway scholarship in order to find instances where the style of Hemingway is given certain properties, while paying special attention to cases where different scholars have given properties to Hemingway's style that do not corroborate each other. These descriptions will then be put through a corpus stylistic exploration of Hemingway's production in order to prove, disprove, or further elaborate on earlier descriptions established by previous research by using modern data analysis tools. This study aims to specify, affirm or elaborate on previous descriptions to

further future research and provide new data to consider when discussing or describing the author's textual properties.

An example of a claim regarding Hemingway's language style comes from Harry Levin's 1957 article "Observations on the Style of Ernest Hemingway" published in *Contexts of Criticism* where he writes: "Hemingway puts his emphasis on nouns because, among other parts of speech, they come closest to things. Stringing them along by means of conjunctions, he approximates the actual flow of experience" (Weeks 79). This claim is also found in the material by Carlos Baker and Jeffrey Meyers, but there it is connected to a temporal variable which means that their claims need to be explored separately from Levin's claim. Hemingway himself also voiced an opinion on his work in regards to brevity, claiming that he cautiously removed anything but the essentials from his writing.

The other category of claims we wish to engage with are those regarding female characters and femininity in Hemingway's work. Gender representation is a large part of both recent and previous Hemingway criticism, and differing descriptions and opinions are found in works by, among others, Edmund Wilson (1940) who noted a growing antagonism towards women in Hemingway's early texts, although one could argue that Wilson was deliberately modest in his wording. As early as 1927, he had defended Hemingway from bad criticism stating that "[t]he reputation of Ernest Hemingway has, in a very short time, reached such proportions that it has already become fashionable to disparage him" (Meyers 113), making him a bit impartial in his assessment. Katherine M. Rogers (1966) and Judith Fetterley (1978) both accused Hemingway of perpetuating sexist stereotypes, Leslie A. Fiedler (1960) argued that he was incapable of creating a female character independent of a man, and Roger Whitlow (1984) believed that his characters reflected his own sexist mindset (Sanderson 171). However, this is not to say that the debate has been entirely one sided throughout the years. Our two research questions based on previous critical material would then be:

- Is Hemingway's textual style constructed as described in the critical material in terms of language patterns and periods of change?
- Can the opinions regarding female representation, character archetypes and participation in Hemingway's fiction be supported by linguistic data?

The first question remains close to previous stylistic endeavors while the second moves further towards literary criticism due to the application of an interpretative theoretical framework. Our hope is that the reader can get acquainted with the corpus stylistic methods and tools while engaging with the first question and then move on to the more interpretative analysis of the second question. There, stylistic method will be combined with feminist theory to show how the stylistic methods can be applied with different aims and goals beyond linguistic data on language style.

#### Stylistics as Method

The relationship between linguistic description and literary appreciation is described by Geoffrey Leech in *Style in Fiction* by using Leo Spitzer's Philological Circle, where Leech writes about a "cyclic motion whereby linguistic observation stimulates or modifies literary insight, and whereby literary insight in its turn stimulates further linguistic observation" (12). This echoes the goals of this study well as it works towards re-evaluating and elaborating on earlier research through new methods.

During the past two decades the line between linguistic method and literary theory has become somewhat blurred by the growing number of studies in the field of corpus stylistics, as shown in "Literary Style and Literary Texts" by Michaela Mahlberg, which describes corpus stylistics as "...the study of literary texts that employs corpus-linguistic methods to support the analysis of textual meanings and the interpretations of texts" (2015 346). Peter



Barry writes about the position of Stylistics in terms of a literary theory in his introductory compilation *Beginning Theory* (3rd ed. 2009), and describes how it is often left out of compilations on the topic of literary theory due to the difficulty of defining it as either a theory or a practice. In this essay corpus stylistics is used for the method, while theory is supplemented where necessary. This means that while quantitative data from the corpus analysis is used to build an argument from, a second theoretical framework is used to create the qualitative analysis. Corpus stylistics lack the ideological and political aspects that are often found in other schools of literary criticism, but in many cases those aspects are added to the interpretation via the additional theory used to understand the results.

The positivist nature of corpus stylistics, in regards to the quantitative data, means that the method claims to achieve quantifiable, repeatable data with a large degree of certainty in accordance with the scientific method, while the qualitative nature of literary theories allows them to claim to represent one perspective of a multifaceted question with a multitude of answers. One could say that the conflict stems from literary theory proposing an interpretation, while corpus stylistics provides observations and numbers. The data must, of course, still be interpreted, which adds a qualitative aspect when the analysis of the data is presented. However, the analysis is often done with an additional theory in mind. The choice of a supplementing qualitative theory also dictates the data selection as it decides which words are chosen as relevant. This echoes Mahlberg's description of corpus stylistics as being literary theory having its interpretation supported by corpus linguistics (2015 346). However, regardless of theoretical pedigree and status, the corpus stylistic approach has provided some noteworthy additions to the scholarship of different authors in recent time.

In 2011, Hannah Spencer published a corpus-linguistic study of H.P Lovecraft's stories using N-grams. N-grams, sometimes referred to as "clusters" or "chunks", are lexical strings that are repeated in the corpus. When performing the queries, one must also specify the

number of collocates to be compared, which is why the N-gram term is preferred as the “N” will be replaced by a numeric in order to specify how many collocates are a part of the repeated lexical string. For instance, tri-gram for strings containing 2 components besides the node or four-grams for a string with 3 components around the node (Crawford & Csomay 54-55). The node word is also included in the count, making the tri-gram string (2 components + node) and the four-gram string (3 components + node).

A good example would be if one imagined a written adaptation of a TV-series, for instance *Friends* or *Star Trek*. In *Friends*, characters are occasionally characterized partially by their repeated speech patterns. The prime example would be Joey, with the phrase “How you doin’?” which has come to define the character and is often used to reference him or his personality by the fanbase or in later spin offs. This would appear as a tri-gram in a corpus query. The phrase also shows how a repeated phrase can function when it is given an additional meaning outside of the words themselves via connotation. In *Star Trek* we find the example “Live long and prosper”, uttered mainly by Leonard Nimoy’s Spock. This would appear as a four-gram in a corpus query and is used a farewell greeting, often connoting a somewhat sad context in the series.

The idea of repeated phrases gaining a further meaning stems from the idiomatic use of language, the idea that language is stored and used in pre-packaged forms that are either used in their basic form or are modified to express a different meaning (Jones & Waller 84). This is further described by John Sinclair, who also coined the term “idiom principle”, in *Corpus, Concordance and Collocation* (1991). Spencer found that an N-gram search in the corpus showed that a certain set of N-grams in Lovecraft’s production were more likely to be found in a negative prosody, or with a negative connotation (Jones & Waller 165). A prosody is a group of words in a repeated structure that gain a meaning in context of each other, in this case a negative description but other examples can be seen earlier from *Friends* and *Star Trek*

where prosodies are used for other functions as well. The use of the negative prosody also appears specific to Lovecraft's language use, providing a new marker of style for the author. This is interesting as it provides new insight into habits of an author that has already received a decent amount of study, and thus also shows how a corpus stylistic study can further elaborate on the style of an author who has already been thoroughly discussed regarding the subject of language style. Spencer's study also elaborates on the earlier concept of something as "Lovecraftian" beyond the archetypes of cosmic horror and mythos that earlier has defined the author's writing towards studies dealing with traits of the Lovecraftian language as well.

Spencer's discovery brings a new item to the table and allows for previous interpretations of mood and environmental descriptions to be re-evaluated in light of a new language habit apparent in the author's work. This makes it possible to identify and discuss hidden discourse in the texts, or understand the subtler descriptive aspects of the texts. A repetition in the pattern regarding descriptions of locations in the author's work also points to setting and environment as interesting topics for further literary analysis, perhaps supplemented by the previous work done on Lovecraft's depiction of ritual and myth, for instance Maurice Lévy's *Lovecraft: A Study in the Fantastic* (1988), in relation to the different locations' function within the story.

Furthermore, Mahlberg's study of Charles Dickens' work in 2013 also provides the discovery of a new pattern, namely word clusters that appear within the author's text in order to create a specific type of characterization (Jones & Waller 164). The study shows that the phrase "with his back to the fire" is commonly used by Dickens to create male characterization in comparison to characterization techniques employed by other authors, similarly to how the previous example from *Friends* now serves to characterize child-like womanizing (Mahlberg 2013 26). From a literary perspective, Mahlberg's discovery sheds further light on a central theme within romanticism. Dickens is by many considered a realist

writer who drew heavily on the earlier romanticism (Fanger, Greiner, Meckier) making a recurring pattern of characterization related to nature, such as a man sitting with his back to the fire, extremely relevant outside of the linguistic interest for a pattern in the author's style. Both of these studies show how corpus-linguistics have been utilized on a large dataset drawn from an author's full production and have resulted in new discoveries about the style of authors who had already been extensively studied and written about previously. The examples deal with sentence structures or phrases, something that has a history of also being incorporated in different fields of literature studies. Due to these points being argued for using corpus material comparisons can also be made to other authors by using reference corpora, making it possible to see if any patterns are specific to the author of the primary corpus, rather than just being traits found in literature in general.

A third example provides a different approach that highlights a major advantage of the stylistic method in comparison to a close reading. O'Halloran's 2007 study of keywords, meaning statistically significant words in comparison to a reference corpus, found in James Joyce's *Eveline* shows how recurring words can further our understanding of underlying general themes in texts. O'Halloran argues that the higher frequency of the modal verb 'would' in relation to the character Eveline shows the theme of her expectations, hypothesizing her future and the possibilities it might provide (Jones & Waller 164). Unlike the two earlier examples, O'Halloran's work is aimed towards finding new information about the character, rather than finding habits and features of style in the author. It also furthers our understanding of the meaning behind the text by being able to base interpretation on one of the smaller components we as readers so often tend to miss. The notion of *Eveline* being a complex text is not a new one, and O'Halloran writes that the subconscious and its function in the text has been discussed earlier, but his study shows that it can be done in a less abstract and more precise manner (2).

This use of corpus linguistics to investigate a detail previously missed as something significant could apply to many general descriptions of Hemingway's production as a whole, especially since the amount of texts produced by Hemingway would make it even more difficult to notice a small detail as something relevant to the author's style as a whole.

Mahlberg points out that Burrows argues for one of the major advantages of corpus based studies being the ability to also put focus on the grammatical words that are often overlooked when assessing or describing an author's production (Mahlberg 2015 351). This is especially important for this project as repeated patterns of representation and participation are central to understanding how gender is created in the texts. The implications of, for instance, a small, repeated verb in relation to a gender signifier could become important for finding any "hidden" patterns of gender characterization.

The descriptions we have decided to focus on in this essay are claims regarding Hemingway's language patterns or tendencies visible through textual habits. In order for a habit or pattern to be considered a feature of "style" it must not only be prominent within the primary source material, but also be a trait specifically found in Hemingway's work when compared to that of other authors. Culpeper and Demmen provide a concise description on the nature of keywords (34-35): "Repetition is the notion underlying both style markers and hence keywords, but not all repetition, only repetition that statistically deviates from the pattern formed by that item in another context" (Culpeper & Demmen 92). This is the reason for the use of reference corpora, as it allows patterns found in Hemingway to be considered traits of his style, not just traits commonly found in language or fiction in general.

The method used in this study varies in practical application, but remains structured in the same way throughout. It could be defined as containing four steps:

**Step 1:** The first task is to identify general descriptions, assumptions and opinions regarding textual style in Hemingway's production, such as his suggested tendency to favour male

characters, his concise sentence structure or repetitive language. This will be followed by a cross-referencing of other claims regarding style, for example the conflicting views on gender representation by Baker and Walsh, working from the same variables. Alternatively, descriptions from other critical or biographical material can be used in order for conflicts of description to be found and defined properly where the above method proves insufficient.

**Step 2:** Moving forward, the proposed feature of Hemingway's style must be formulated into a query suitable for whichever tool deemed most likely to produce results that corroborate, disprove or elaborate upon the claim. In essence, this step is concerned with creating a query that can provide data relevant to the claim we intend to investigate. The results can then be interpreted in a way that answers or elaborates further upon the question. The idea is to interpret the resulting data in a manner that creates a nuanced and satisfying conclusion, without crossing the line to pure speculation. Some interpretation is necessary as "the study of the relation between linguistic form and literary function cannot be reduced to mechanical objectivity" (Leech 3).

**Step 3:** At this step, the data from step two is viewed in context of the initial claim to see if the data supports the interpretation of it, disproves it, or provides new insight into the actual features and functions that create the basis for the claim in regards to Hemingway's style. Then the defined feature will be compared to results from one or several reference corpora in order to see if the feature is truly specific to Hemingway, and not just to literature of the 1900's in general.

**Step 4:** Finally, the findings are summarized in a way that enables further understanding of an authoritative claim about the author's style. These findings are then put in relation to other claims regarding the same period, feature or tendency.

However, some claims need several queries in order to be properly represented. For instance, when looking at female representation an adjective could be found in different positions in

relation to the noun, making it necessary to repeat step two and formulate several queries for the same question.

### Feminist Literary Criticism

For the questions regarding female representation and participation, a framework of interpretive theory is added, and the results are analyzed from a perspective based in feminist literary criticism. The literary theory is used to formulate the queries, focusing on verbs for participation and adjectives for representation, and the query is subsequently created to provide data on the topic. The results are then interpreted qualitatively by applying the theoretical framework to the data. This is in line with Mahlberg's previously cited description of how corpus stylistics should be used in practice.

The situation for women during Hemingway's lifetime was subject to radical change. By the time of release for his first collection of fiction *My Old Man* in 1922, a "new" type of woman had emerged in the American society. According to James Nagel, women were now publicly smoking, taking part in parties with heavy drinking, as well as being able to divorce men who trapped them in unfavorable marriages (MacDonald 92). Furthermore, the American suffragette movement succeeded in passing the Nineteenth Amendment, which enabled women to vote, on August 26th 1920 and granted women a drastically more equal place in society (Hotchner 42). While the feminist movement had begun gaining serious momentum in America during the 1920's, it had progressed even further in Europe and most notably in Paris. As Hemingway lived in Paris between 1922-1928 and actively took part in circles that embraced the political and social changes for women, he was undoubtedly influenced by these changes mainly through Gertrude Stein's tutelage and friendship.

Stein was an embodiment of the new wave of feminism which, in the words of Hotchner, "bobbed her hair, smoked continually, walked about Paris unaccompanied, and bartered with sign makers, window washers, and booksellers, all of which established her as

one of the first liberated women of the century” (66). A character with several similar traits is found already in Hemingway’s second novel *The Sun Also Rises* (1926) in the character of Brett Ashley who was both sexually liberated as well as free-thinking, independent female (MacDonald 92).

Towards the end of Hemingway’s life, a new wave of feminism had started to emerge. Simone De Beauvoir’s 1949 *The Second Sex* shows, among other things, the importance of female representation in fiction as it argues that femininity is not inherent at birth, but is created through external factors, as “one is not born a woman, but, rather, becomes one” (12). The feminine identity is, according to De Beauvoir, historically created as subservient in relation to the masculine, and is taught to young women until it is internalized and perpetuated by the individuals themselves. In relation to the topic of this essay, De Beauvoir’s explanation of how femininity is created shows the importance of acknowledging female representation in the novels as a powerful tool for creating characterization.

Whereas previous incarnations of the ideals of equality had been focused on social interactions, identity and public behavior, the 1950’s and 60’s came to see a movement towards women’s rights in the academic world and the workplace. The importance of meaningful labor as an important part of enabling women to find their identity and lead purposeful lives is, amongst other places, written about in *The Feminine Mystique* by Betty Friedan. Friedan’s book became a bestseller when it was published in 1963, proving that the ideas described in it were alive and likely a part of the discourse at the time of Hemingway’s final publications.

The concept of activity as an important part of creating, and describing, identity is combined with De Beauvoir’s emphasis on social engineering by Judith Butler in *Gender Trouble* (1999). In the text, Butler discusses the necessity of the identity labeled as “woman” for feminist politics, but also problematizes the exclusionary element of such an identity (9).



This is relevant for this study as we do, indeed, assume the existence of a stable group related to the identity “woman” in Hemingway’s text through the method of engaging his texts through queries based on gendered nouns and pronouns. However, Butler points to this as something inherent to the nature of language, and since language is the material which is investigated here the assumption is a necessary one. Butler partially covers the same ground as De Beauvoir, but with an added elaboration on gender identity beyond the binary (10). *The Second Sex* argues for the separation of the concepts of sex and gender, the former being the biological body and the latter the role or identity. When we perform our queries we will, as previously mentioned, use words coded from the gender dichotomy as the starting point due to these reliably marking the presence of a gendered character in the text. Yet, one must acknowledge that these words are coded on the basis of a binary, heteronormative sex as experienced by the narrator.

While this goes against De Beauvoir’s perspective on the separation between sex, as biology, and gender as a fluid identity, the actions or descriptions highlighted by the queries should still be seen with this separation in mind. In fact, the way this essay presents the data is closer to the relation between sex and gender argued for by Butler, that they are both constructed, and neither is an inherent feature of biology (11). The nouns might exist in an exclusive dichotomy, but the adjectives and adverbs do not. Butler also covers the use of activities linked to gender identity as a way of “performing” our gender roles on a daily basis: “[g]ender is the repeated stylization of the body, a set of repeated acts within a highly rigid regulatory frame that congeal over time to produce the appearance of substance, of a natural sort of being” (43–4). Because of this the activities found in relation to our queries should be seen as creating a gender identity, not defining a pre-existing category of biological sex.

We will start by looking at how the female characters in Hemingway’s fiction “perform” their roles as women by looking at what actions they are most commonly

undertaking. By looking at gender participation in the texts we will also be able to see how these actions show the relation between male and female characters in the stories. How women are depicted, and represented, will be explored through an analysis of how women are described by Hemingway's male characters, thus showing how the author represents women in the texts. The narration being handled mainly by male characters echoes De Beauvoir's statement that "[r]epresentation of the world, like the world itself, is the work of men; they describe it from their own point of view, which they confuse with absolute truth"(196).

In order to avoid viewing the quantitative data as 'the absolute truth' of the representation of women, it is important to note how the male gaze functions in the stories as well. This is especially noteworthy since most of Hemingway's narrative characters are male, as shown in the later section on gender participation. The preference for male characters means that this essay will not only show how women are described in general in

Hemingway's work, but specifically how they are described from a male perspective. Both of these categories must be seen from a perspective that compares the data for both gender groups, as a description or action that is equally represented in both male and female characters cannot be considered simply masculine or feminine, but rather as a signifier of a different group of characterizing features.

In terms of the method for this essay, female participation and representation will be looked at separately in order to investigate both how female characters are described and which attributes are ascribed to them by looking at female representation, as well as exploring which actions are performed by the female characters in order to see how the female gender is performed in Hemingway's fiction through female participation in the text.

### The Specialized Corpus

In *Doing Corpus Linguistics* (2016) by William Crawford and Eniko Csomay, the topic of building a corpus is discussed quite extensively. The authors state that the efficiency of a

personal corpus is highly dependent on a clear and concise statement of purpose, a research question or topic (78). A clear statement allows for a smaller curation of materials for the corpus, but one must still consider the fact that a larger amount of text makes for data that could be considered representative to a higher degree. The purpose is to look at features in fiction written by Hemingway, and the corpus needs to contain as large a part of his production as possible. The closer the corpus is to containing the entirety of his fictional output, the closer it will be able to capture repeated patterns in Hemingway's writing. In order to be considered representational the corpus should also contain, and be able to differentiate between, the different types of fictional outputs where Hemingway was published, such as novels, magazines and different types of collections.

The format of the material must also be taken into consideration, and we decided that this corpus would use .txt-files due to the simplicity of compiling them and the general simplicity of handling them in regard to software compatibility and storage. The availability of the desired material made the initial curation quite simple, but Crawford and Csomay also stress the importance of taking copyright laws into consideration when making a corpus public (76). Due to this, the primary Hemingway corpus used for this study is not publicly available, but the authors welcome any questions regarding further use of it. Two texts from Hemingway's bibliography, *The Torrents of Spring* (1926) and *To Have and Have Not* (1937), had to be omitted due to issues obtaining them in the correct format. The structure of the Hemingway corpus is based on the temporal aspect. This is also the primary factor when creating sub-corpora within the main compilation. The sub-corpora are formed by compiling texts in accordance with the periods of Hemingway's productivity provided by Hotchner in *Hemingway and His World* (1990). As suggested by Crawford and Csomay, the individual publications have been kept separate in order to make it possible to compare specific texts to the rest of the corpus, or a reference corpus (82). There are also sub-corpora for each of

Hotchner's periods, where the texts published during the relevant time period have been collected into one file. Hotchner's periods are as follow:

The Early Years: 1889-1921

The Paris Years: 1922-1928

The Key West Years: 1929-1936

The Spanish Years: 1937-1940

WW2 And Hemingway: 1941-1944

The Cuban/Venetian Years: 1945-1953

The Dangerous Years: 1951-1961

As the titles used by Hotchner might suggest, these periods are defined by the specific circumstances during which Hemingway was active, or the locations he was active in. By using these sub-corpora this project will be able to view proposed stylistic features with a temporal or spatial variable. Not all periods are represented in the corpus, as some of them are used strictly to describe a personal period of the author in a biographical sense, without any publications taking place.

The observations will mainly be made through the use of Laurence Anthony's freeware concordance tool AntConc created at Waseda University in Japan. However, Mike Scott's program WordSmith v.4 will also be used for sentence length and type/token-ratio due to these statistics being included in the tool's statistics tab without requiring further calculations on our part. The type/token-density tool can be used to look at traits in Hemingway's texts that deal with density of his writing in terms of word choices or the repetition of phrases and will be further elaborated on in a later section.

### Source Selection

The overlying criteria for the selected main critical material within the essay has first and foremost been the frequency of which the source has been cited within other works, proving its legitimacy among the scholarly community. Furthermore, as the subject of various Hemingway criticisms has been undertaken and evolved for more than fifty years, even dated sources may prove valuable if they have remained relevant despite the test of time and intense peer scrutiny. The information necessary for the questions posted in the essay hinges on these points, which are closely interlinked with the previously stated research questions:

- Finding scholarly or autobiographical works which handle the subject of Hemingway's literary style connected to any change in external circumstances surrounding Hemingway's personal growth as an author.
- Analyzing critical texts which handles Hemingway's use of female characters versus male, to acquire an understanding of both the current as well as the past state of the discussion on the subject.
- Through empirical studies determine the level of accuracy and/or relevancy of the collected data in relation to the selected method of the essay.

For the first level of the study, the main focus was put on finding which authors held the most sway in the academic community in order to take part of what is considered seminal works on Hemingway himself. By utilizing the Google Scholar search engine to enquire about the frequency of certain authors being cited, two sources providing relevant background material stand out regarding number of times cited, and the generally positive attitude towards them. These two were Carlos Baker's *Hemingway: The Writer as Artist* (1998) and Jeffrey Meyer's *Hemingway: A Biography* (1999), which are the two most cited texts on Hemingway

according to the search engine. For his book, Carlos Baker was granted access to hundreds of Hemingway's manuscripts and personal letters shortly after the author's death, and spoke to a large number of people who knew him personally as well. The result is a book that judges Hemingway's works through a detailed critical study which is rich in background information and biographical material, some of which has been added through revisions at a later point, as the original was published in 1952, years before Hemingway died and his posthumous works were published. It provided ample research materials concerning Hemingway's many influences and different periods of writing, as well as how they shaped him as an author.

While Baker sheds light on the values and influences of Hemingway's texts, Jeffrey Meyers goes for a more academic summary of both the private and the professional sphere that describes rather than evaluates. Some other frequently cited sources in this essay come from *The Cambridge Companion to Ernest Hemingway* (1996) edited by Scott Donaldson, author of *By Force of Will: The Life and Art of Ernest Hemingway* (1977). Donaldson has written extensively on Hemingway and many of his contemporaries, and we consider his selection of texts authoritative. In *The Cambridge Companion to Ernest Hemingway*, he has amassed a selection of the most recent critical endeavors on the author himself as well as his works.

The chapter most suited for our research on Hemingway's influences is the contribution made by Elizabeth Dewberry, called "Hemingway's Journalism and the Realist Dilemma", which focuses mainly on his early years in the journalism business, and offers valuable information on the early influences of his and how they permanently shaped his future fiction and nonfiction alike. Furthermore, Rena Sanderson's "Hemingway and Gender History" provides ample material for the discussion regarding the history of the debate concerning Hemingway's supposed misogynistic tendencies. Another useful part of the source

is the selected bibliography, which provides more material on other topics that are related to our area of research.

Many other sources have been selected for the various parts of the essay, for the sake of demonstrating the different perspectives in Hemingway criticism, for instance voices on gender cited in the relevant additions to *The Cambridge Companion*, have been used and engaged with in order to define opinions on Hemingway in regards to the topic of gender, and sources dealing with either Hemingway's style or his style in relation to criticism have been used to create the background for our research on the author's language habits. Hemingway's language style is described by Levin as putting the emphasis on nouns. This is elaborated on by Meyers and Baker, who give a date for this trait appearing in the author's writing.

The topic of gender is written about by Wilson in terms of a growing antagonism towards women, while Rogers and Fetterley state that Hemingway's female characters are sexist stereotypes. Furthermore, Fiedler views Hemingway as incapable of creating a female character independent from their male co-characters. Finally, Whitlow argues that these tendencies in Hemingway's female characters reflected the author's sexist mindset (Sanderson 171). Based on this material, there are many interesting aspects to search for.

#### Modes of Enquiring the Hemingway Corpus

The method will employ four different main categories of queries in order to find and define the language patterns used by Hemingway as suggested by the critical material. Before any queries can be made, a word list must be compiled. The general word list contains the entirety of the available text material, and is, in essence, the words found in the corpus sorted by frequency of occurrence. The list does not contain every single instance of a word appearing in the text separately, but registers every word as a "Type" with a noted frequency of how many times the type appears in the text. When referring to words on a separate level, as in

every instance of occurrence when providing the word count for a corpus or section, the term “Token” is used.

The primary corpus of this project is specified in appendix A and contains 69 texts, but it could be generalized as containing the 62 available fictional texts produced by the author Ernest Hemingway and published between 1920 and 1960, while also containing some material written by him during that period but published after his death in 1961. The Hemingway corpus includes both novels and short stories, as well as the unfinished stories or proto-manuscripts published after his death and has a word count of 567,596 tokens. The reference corpora for this study are the COHA (The Corpus of Historical American English) mixed material sample corpus of 3.6 million words, which includes different types of texts and transcriptions, the COHA Fiction literature corpus that contains literary texts from 1900 to 1960 and has a word count of 69,368,318, license being provided by the Linnaeus University in Växjö, and finally the Corpus of English Novels, which was compiled, and provided to us, by Henrik De Smet at University of Leuven and has a word count of 54,871,679.

The reason for using three different reference corpora is that they relate to different periods or types of text. For instance, the COHA Fiction is used to give comparative values for fiction literature published between 1900 and 1950 and the COHA sample corpus allow for comparisons with general language of different types. De Smet’s Corpus of English Novels contains material dating from the century before Hemingway, which provides a perspective on what types of language was used in novels during the time before Hemingway started writing. Having more material to compare to enables this project to take more variables into account when attempting to make sense of the results.

The next mode is the use of a “Keyword” list. Keywords are not found by a simple parallel comparison of token data between corpora, but takes other variables into



consideration as well, depending on the statistical method used to create the keyword list. This study has used a log-likelihood calculation built into AntConc's keyword list generator, which is one of the more commonly used statistics for significance when comparing corpora of different sizes. Log-likelihood calculates statistical significance of occurrences based on number of hits and corpus size, and then creates a list based on a "keyness" rating which rates the comparative statistical significance of the usage in the primary corpora in relation to the reference corpus. A higher "keyness" rating indicates a higher chance of the difference being notable in the corpora even when taking the different size of the corpora into account, thus making the result statistically significant.

The list of keywords can then be utilized in several ways. Different data can be selected from the list of keywords when dealing with claims on style as heightened frequencies might indicate new style markers or specific traits of the author's language habits (Leech 56). Much of the content found in both the general wordlist and keyword list will likely be irrelevant due to "a reasonable assumption that even the most ingenious author will not use every part of the linguistic code for particular artistic purposes" (Leech 56), but the keyword list makes it possible to identify and explore patterns and phenomena specific to the primary corpus. This is especially true for this study, where the query must also be justified by a claim found in the critical material. When claims are found where the keyword list can be used it is important to note that focus will not be on the singular word type suggested by the claim, but also the textual setting and environment. This will take the shape of a Keyword in Context query (henceforth KWIC), which looks at the immediate surroundings of the word type and presents it in the form of a concordance line (Crawford & Csomay 36-37).

A concordance line is a line of text containing the keyword, enabling the viewer to read the word or sentence in its original environment. Concordance lines allow for an overview where general patterns can be discovered. Text type specific tendencies become

visible and can either help define the specifics of a certain type of text, for instance ‘may’ and ‘must’ in legislation as shown by Jones & Waller, or the style markers of an author (74). This project will look for patterns of collocation and colligation, the former being the relationship between the keyword and individual words and the latter being the relationship between the keyword and grammatical categories. These two patterns allow for discoveries similar to O’Halloran’s ‘would’ in Joyce’s text. For example, if the colligation of a masculine pronoun would commonly be a verb it would indicate action as a concept the author connects to masculinity, and if a common collocate of the same pronoun would be a specific verb this would further specify the type of action favored by the author when writing male characters.

The next mode of enquiry will be N-gram queries, utilizing the list of keywords as a starting point. N-grams are, as mentioned in the section on method regarding Spencer’s study on Lovecraft, lexical strings that are repeated in the corpus. The importance of the idiomatic phrase as a marker of style is found in Mahlberg’s data on Dickens’ characterization of male characters as well, where a lexical string recurs as the author’s pre-packaged phrase for masculine depiction, thus creating a stylistic trait. By performing the N-gram queries, with the use of a node word being the only specific component, this project intends to find similar habits of expression in Hemingway’s work.

The third mode of enquiry will be the search for semantic prosodies. Semantic prosodies are “a form of meaning which is established through the proximity of a consistent series of collocates” (Louw 57). This project will mainly focus on affective meaning created by repeated word combinations collocating to the node words from the list of keywords. Where the keywords and N-grams can show how Hemingway designed certain traits of his work in terms of language, the prosodies will help provide insight into the attitude or emotions towards the concepts as they are depicted by language patterns specific to the author. In combination the keywords and N-grams can show a pattern, and the prosodies can

explain what the author intended to emote by using it. According to Louw, the method of looking at semantic prosodies is primarily useful when attempting to gain insight into the attitude or evaluative opinions of the writer, or speaker, regarding a certain subject or keyword (58). This is also the way semantic prosodies were utilized by Spencer when looking at descriptions of locations in Lovecraft's work.

Finally, the fourth general mode of enquiry at our disposal is the use of Part-of-Speech tagging software to look at grammatical functions and tendencies in Hemingway's works (henceforth PoS-tagging). PoS-tagging is a complex endeavor when one is tagging and handling large amounts of tag types, but there are always more simple enquiries available to facilitate the process. "Tag types" is the number of specific tags the software can use to mark individual words in a text. On a basic level, tagging could utilize a small selection of categories, such as only nouns, pronouns, adverbs, adjectives, conjunctions and prepositions, but it could potentially become infinitely more complex as further categories are added. This is especially true for words that gain their role or function from context. This study will use the TagAnt tool, which is provided by the same Laurence Anthony that created AntConc, for tagging the primary corpus. TagAnt is a streamlined user front for the TreeTagger software created by Helmut Schmid at the Institute for Computational Linguistics of the University of Stuttgart. Schmid's system is capable of defining 58 different tags in a text, which should cover the needs of this study rather well.<sup>1</sup> The PoS-tagger will be used to tag the corpus according to parts of speech, enabling AntConc to look at specific grammatical structures or the nature of a specific grammatical function in relation to a word class through using the tags in different queries. For instance, adjectives collocating to a specific noun could provide

---

<sup>1</sup> Schmid's "Improvements in Part-of-speech tagging with an application to German" (1995) shows the success rate of the tagger to be above 90%.

insight into how that noun is represented in the texts, or a verb collocating to a pronoun could elaborate on the common actions of a character type connected to the pronoun.

### Hemingway's Background and Influences

Initially, Hemingway's writing style was mainly derived from the classic rules and traditions of journalism, up until the 13 years long break from the medium between November 1924 and his 1937 North American Newspaper Alliance (NANA) coverage of the Spanish Civil War (Baker 14-16). During Hemingway's break with journalism, with the intent to pursue new literary techniques, he spent time in Paris under the influence of prominent names within the emerging modernist movement such as Gertrude Stein, and other American expatriates including Ford Madox Ford, F. Scott Fitzgerald, Sherwood Anderson and Ezra Pound. By 1923, Hemingway had worked as a reporter for the *Toronto Star* on several occasions while being trained as a cub reporter for the *Kansas City Star*, all this while also writing for his high school newspaper. It was here that he made out his first set of rules for the art of writing (Donaldson 16). Between 1920 and 1960, Hemingway wrote for some of the great American newspapers, ranging from articles in lifestyle magazines such as *Life*, *Look* and *Esquire*, to war journalism during the Spanish Civil War and WWII England (Meyers 51). His position at the *Star* enabled him to choose freely between both subjects and styles that suited him the most in his forthcoming career within literature, which had a lasting impact on his prose. As Donaldson (19) has observed, the *Star* supplied him with a standardized stylesheet of 110 rules, which taught him the fundamentals of his prose: short sentences and first paragraphs, avoidance of adjectives and superfluous materials, and minimalistic yet rustic English were all things that stayed with Hemingway practically his entire life according to the critical material.

Arguably, anything produced either between 1922-1928 or shortly after, moves further and further towards the modernist spectrum of style as Hemingway began to break away from

his journalistic roots, all while maintaining the same distinctions between journalism and fiction, real and imaginative, that he always had (Hotchner 50-1). Jeffrey Meyers also writes about Hemingway's influences at the start of his career in writing serious fiction somewhere between 1922-23 being Stendhal, Flaubert, Maupassant, Twain, Conrad, Stein, Pound and Eliot within the American tradition and Tolstoy, Kipling, Joyce, D.H Lawrence and at a later stage also T.E Lawrence internationally (72). He was also a close friend to Joyce, Ford and Eliot, while also being a student of Ezra Pound.

According to Dewberry (25), Hemingway's devotion to the rules of journalism is most evident in these early years, as he used his time with the *Star* to flesh out the framework for his later literature by separating the necessities from the redundant. Hemingway was adequate at writing articles and other correspondence, yet he saw it only as a mean towards an end. In a 1924 letter to the *Transatlantic Review*, he stated that "the only reason for writing journalism" was to get paid well, as "when you destroy the valuable things you have by writing about them, you want to get big money for it" (Baker 9). Travelling wherever he thought he could find a good journalistic scoop or idea for a story, he visited the guerilla-infested backstreets of Genoa in 1922 as well as Greece in shambles at the end of the Greco-Turkish war that same year. Through his journeys, Hemingway educated himself within the subjects that would stay within his writings for a large amount of his bibliography: man versus nature, simple pleasures such as food and travelling, sports, politics and war (Donaldson 23).

The combination of journalistic beginnings, artistic influence through other authors and experiences through travelling all led to a heightened state of realism within his early works of fiction, as much of the material uncovered for newspaper articles bled into his stories. For Hemingway's first published full-length work of fiction *In Our Time* (1924), no less than twenty-five of his journalistic articles from his time at the *Toronto Star* shared similarities with events described in the novel, in an attempt to ground his fiction in reality.

This was done in order to communicate to his readers some basic, universal truths about the world without going into too much individual detail. The process, which he called the “iceberg theory”, stated that “you could omit anything if you knew that you omitted and the omitted part would strengthen the story and make people feel something more than they understood” (Hemingway *75 Feast*). It was another remnant of a rule he picked up during his early journalism years which remained throughout his life.

The task of providing a general description of Hemingway’s style through understanding the author’s influences is further complicated by Hemingway’s own reactions regarding other people interpreting, defining or criticizing his work: “They’re the people who hear an echo and think they originated the sound” he once told his brother Leicester, addressing those who thought of him as a liar and a fake, “They hear or read somewhere I’m a phony and it’s suddenly a fact in their minds. Like Heywood Broun branding me a phony on boxing... I’m getting plenty sick of this branding, and it probably hasn’t even run its course... the only way I’m a phony is in the sense that every writer of fiction is: I make things up so they’ll seem real” (L. Hemingway 161). While this quotation might not directly deal with criticism of his textual style, it shows the difficulties involved in maintaining a dialogue between the author and his critics, even those fortunate enough to have been alive during the same time as the author. This also necessitates a thorough review of Hemingway’s own opinions of his style, as they might very well present a differing opinion when compared to his critics. Hemingway’s own descriptions of his style will be considered of the same validity as the opinions of a scholarly critique of the text, neither more nor less. This is due to the absence of the author in the reading process as described by Roland Barthes in his 1967 essay “The Death of the Author”. Barthes states that once a text is finalized by the author, the creator of the piece loses control and the interpretation is carried out on an individual level by the reader. One can, and many do, attempt to discern authorial intent when interacting with a

text, but the reading is still shaped by the reader's personal background knowledge and experiences. Furthermore, Barthes writes about how the text is not the work of one single author, but rather a large amount of previous stories, cultural content and pre-existing materials brought together through one person but created by society as a whole. Due to this, Hemingway will be allowed place to argue for his own reading and interpretation of his works in the same way as other critics, rather than being granted a separate place as the primary source of authorial intent.

Being influenced by the literary movement named "The Lost Generation", which was in constant change and under the influence of varied community of writers, Hemingway is difficult to generalize or attempt to describe in terms of textual style without either leaving something out or disregarding some features in order to highlight others as prolific. Scholars such as Jeffrey Meyers, Carlos Baker, Sheridan Baker and Scott Donaldson offer critically acclaimed pieces on both Hemingway's works and his life. However, none of these works utilize a quantitative method to further specify and elaborate upon the descriptions of Hemingway's text, thus perhaps missing nuances in language use easily glanced over when conducting close readings of the works.

#### Hemingway's Style in Data

On the topic of style, adjectives and nouns are central components to claims connected to Hemingway's tendencies towards short language structures and few words. This connection is, for instance, made by Harry Levin in his 1957 article "Observations on the Style of Ernest Hemingway" published in *Contexts of Criticism* where he writes: "Hemingway puts his emphasis on nouns because, among other parts of speech, they come closest to things. Stringing them along by means of conjunctions, he approximates the actual flow of experience." This claim is also found in material by both Carlos Baker and Jeffrey Meyers, but there it is connected to a temporal variable, and will be discussed with its own set of

specifics in a later section. As for exploring Levin’s general claim, the method used will be to create a list which shows the different function tags used by TagAnt in order according to frequency. This will provide data not only on the use of nouns in relation to adjectives, but also on other language functions that could elaborate on the idea. An overview of adjectives and nouns is seen in figure 1 below.

<b>Corpus:</b>	<b>Hemingway</b>	<b>CEN</b>	<b>COHA Fiction</b>
Nouns/ 100k	20,237	10,426	24,294
Adjectives/ 100k	4,948	2,947	5,682
Size in tokens	567,596	54,871,679	69,368,318

*Fig 1: Overview of noun and adjective use in Hemingway and reference corpora. Hemingway is shown to be between the CEN and the COHA, while closer to the COHA.*

When investigating Levine’s claim, we will use the PoS-tagger TagAnt and the concordance tool AntConc. TagAnt will be used to create tagged versions of the different corpora, enabling searches on specific parts of speech to be performed. The base tag for the different groups of nouns is “\*\_N\*”, which can then be further modified to specify singular, plural and proper nouns. The “\*” signifies a “wildcard” search, which means that AntConc will count every occurrence regardless of what is found in that position and thus include different types and noun categories in the total count.

A wildcard search on the noun base tag from the PoS-tagger gives every occurrence of a noun in either singular or plural. The query provides 98,374 hits in the Hemingway corpus, giving a standard frequency value of 17,332/100,000. Adding to this number, proper nouns in singular and plural appears at a standard frequency of 2,905/100,000. This means that the ratio of nouns in Hemingway’s texts would be 20,237/100,000. In order to understand whether or not these numbers are significantly high or low, comparisons must be made to the

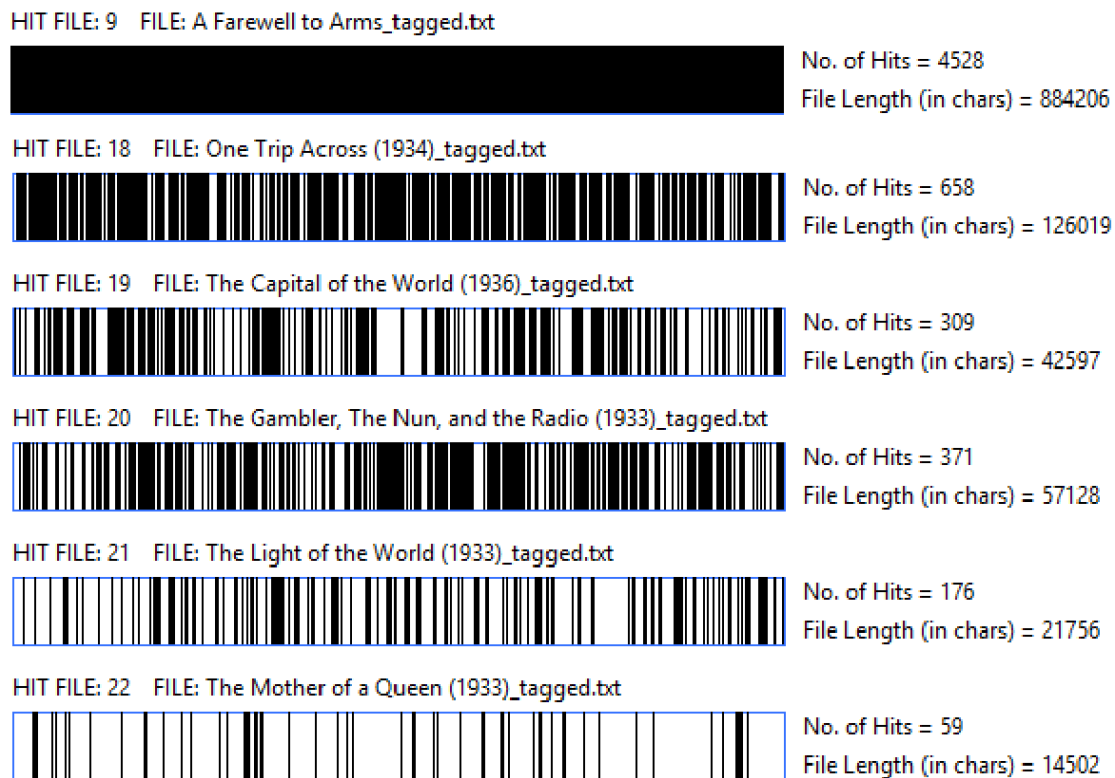


different reference corpora. This will also make it possible to see how the results for Hemingway differ to different types of language and different times of production. For instance, Hemingway could be shown to use language more similar to spoken language, or novels from either the 1800's or his contemporaries of the early 1900's.

In comparison, the COHA Fiction reference corpus gave 12,589,097 hits for the query on singular and plural nouns, providing a standardized number of 181,481.9/1,000,000 and 4,263,924 hits, giving a standard frequency of 61,467.8/1,000,000, for proper nouns. This gives a total noun frequency of 242,948.9/1,000,000 which translates to 24,294.89/100,000. The CEN corpus gives 4,554,841 hits for nouns in plural or singular, and 1,166,146 for proper nouns for a total count of 5,720,987. The adjusted standard frequency for the result is 10,426 /100,000. A similar query is made for adjectives using the adjective base tag, which provides 28,089 hits for adjectives. This gives a standardized frequency of 4,948.7/100,000. The CEN gives a standard frequency of 2,947.3/100,000 and COHA Fiction produces 5,682.4/100,000 for adjectives in the reference corpora. Hemingway's use of nouns and adjectives places him between the reference corpora, although closer to his contemporaries in the COHA Fiction than the previous period represented by the CEN.

Before attempting to interpret the results, one must be keenly aware of the corpora used and how their different attributes could have influenced the results seen in figure 1. The Hemingway corpus is the smallest by far, with both reference corpora being more than a hundred times larger. This offers a better representability of the literary periods for both the reference corpora, 1800's for the CEN and the first half of the 1900's for COHA Fiction, but it also dilutes any author specific traits, or styles, that could possibly have been more fruitful to compare with Hemingway directly. A move towards a heightened use of nouns and adjectives is seen from the century represented in CEN compared to both Hemingway and his contemporary COHA Fiction.

Hemingway follows this trend in his use of both nouns and adjectives, but is still found at a lower frequency in both cases compared to the average 1900’s author as defined by the results from COHA Fiction in both categories. Hemingway’s use of nouns is on par with the “industry standard”, but a very interesting picture is painted when looking at adjective use. One would expect language use to be evenly split in terms of parts of speech use or changes to be confined to some specific period of time due to a person’s language style being coherent and changes taking time gradually over time. Hemingway, however, varies greatly in his adjective use throughout his career and even during decades (see figure 2).



*Fig 2: Adjective density in Hemingway’s work between 1929 and 1936. Adjectives in the text are marked in black.*

As seen in figure 2, the adjective use varies greatly and a similarly spread out pattern is found in all of Hemingway’s periods of production. In fact, most of the author’s adjectives seem confined to a certain set of works while the majority conforms to Levin’s description. The example of high adjective density shown in figure 2 also happens to be a novel, but a

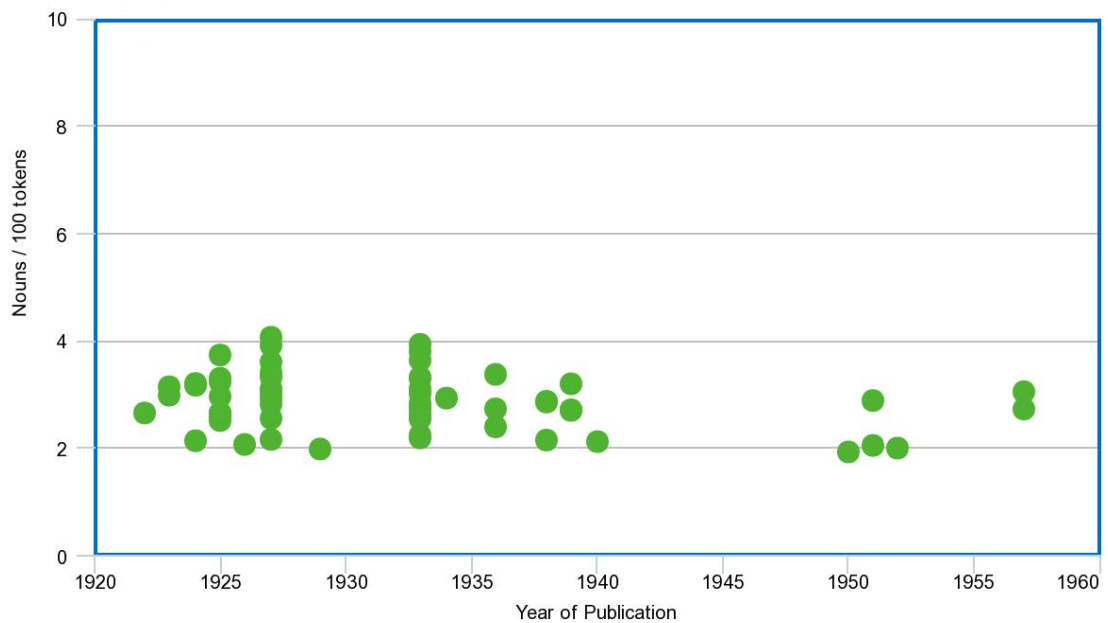
similar density is found in the short stories “One Trip Across” and “Night Before Battle”. The novels do, however, use more adjectives and are found at a higher density of usage throughout. It is important to note that the novels generally do appear much denser when plotted next to the short stories due to their larger size in terms of tokens, so side by side comparisons must be viewed with the numbers presented on the right hand side of the figure in mind rather than relying only on the visual representation. The shorter texts can be compared visually, but the longer novels become fully marked due to the size of the plot. Because of this the novel has been left out of figure 3, but a similar issue appears in the plot for “One Trip Across”. Returning to Levin’s description, nouns are found at a much more even rate throughout the corpus as shown in figure 3.



*Fig 3: Noun density in Hemingway 1933 to 1936.*

This relationship contributes towards a possible explanation of the descriptions of Hemingway’s work as being heavy in the usage of nouns even when that is not seen in the general frequency comparisons. The plots indicate that Hemingway’s use of adjectives is mainly found in a few of his texts, while the use of nouns remains more constant throughout. This is seen in figure 4, where the results for individual texts are shown standardized per 100

words. Please note that the adjectives are provided in scale 0 - 1.2 while the nouns are given in 1 - 10. This is due to the different sizes of the results in the different categories. The larger fluctuation of frequency found in adjectives enables the texts to rely more heavily on nouns in general terms, but this fact would then be hidden behind the high adjective density found in a small number of texts when comparing the totals. It is interesting to note that the segment of texts containing fewer adjectives disappears after 1940, hinting towards an evolution of the author's adjective usage. However, as seen in both the total numbers given above and the visual representation in figure 4, Hemingway indeed prefers nouns to adjectives.





in massive proportions, I polish into a tiny gem. (Hemingway in *Playboy Magazine*, 1963:10:01)

Hemingway's short structures and decisive use of language are often mentioned in the general descriptions used to define his writing, and this notion will be investigated in this section, but Meyers also provides a year for when this style marker became a habit for the author. Hemingway developed an "obsession with counting words" and repeating phrases around the same time that he began referring to himself under the moniker "Papa", which he first started doing in notes during July of 1934 (240). The starting point here should be to validate the notion of short sentence structures in Hemingway's writing, and then moving on to see how they might have changed during his career, especially after 1934. After exploring the sentence length this study will move on to look at variation and repetition in word use through comparing type/token density in the different corpora.

WordSmith's statistics tab offers a mean length of sentence structure statistic within a corpus, and the data provided for the entirety of the Hemingway corpus is a mean word count of 10 words per sentence structure. This might seem perfectly normal, but the COHA reference corpora provides a mean value of 17 for the mixed material, and 13 for fiction between 1900 and 1960. This means that that the Hemingway corpus contains a shorter sentence length in comparison to both, especially to the mixed material corpus. The CEN corpus gives an average sentence length of 16, creating a perspective from where Hemingway's language appears very much as described in the critical material in terms of comparative sentence length. When performing the same queries on the individual sub-corpora, a more interesting picture appears in connection to Meyers' claim of changes taking place after 1934.

In the period between 1922 and 1928, which Hotchner refers to as The Paris Years, Hemingway’s output varied greatly in sentence length, but where still short in comparison to both later periods and reference corpora as visualized in figure 5 below.

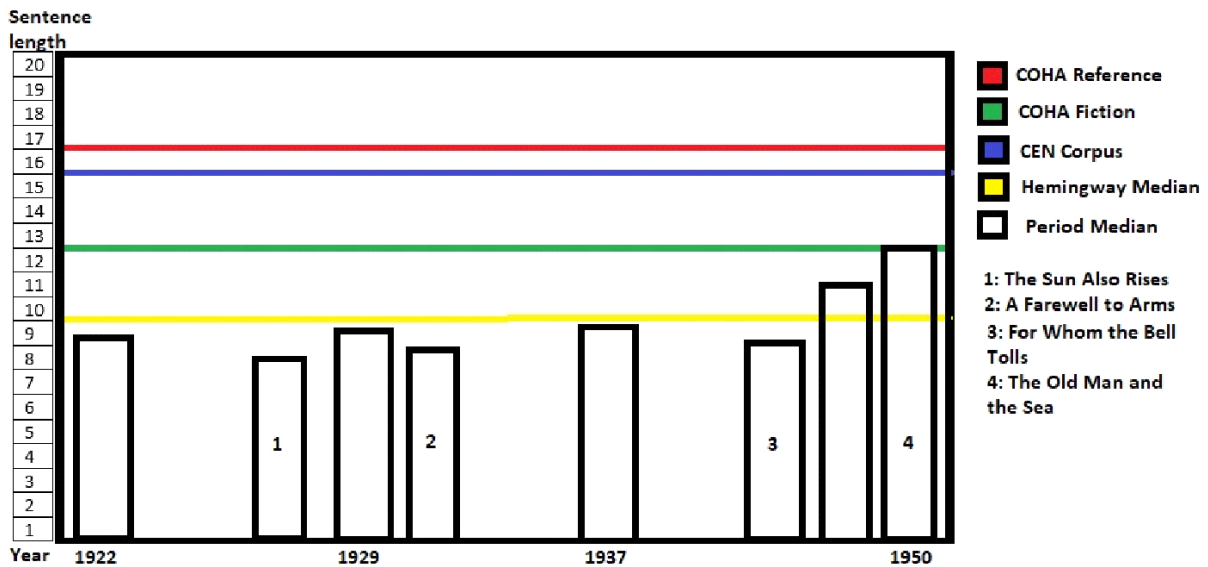


Fig 5: Comparative overview of sentence lengths over time.

The fluctuations are most noticeable in the short stories, although still present in the novels, and an explanation for this according to Hotchner is that during this time Hemingway was surrounded by highly influential authors and artists who inspired and supported him (60). It was during this time that he abandoned journalism at the behest of Gertrude Stein, and began to seriously develop his prose (C. Baker 117-119). Commenting that “[i]f you keep on doing newspaper work you will never see things, you will only see words and that will not do” (Hotchner 60). Stein’s influence over Hemingway was so strong that when he showed her his first draft of *The Sun Also Rises*, she told him to start over again and concentrate harder, which he ultimately did (Hotchner 60).

In this period, one finds “My Old Man” (1922, ed. 23), which boasts the highest sentence length of the entire Hemingway corpus at 25.5 words per structure, but also “The Killers” from 1927 which has a low value, both in terms of the reference corpora and Hemingway, of 7.4. The period is very broad in this aspect, and middle-grounds on both sides

of the spectrum were published. “In Another Country” from 1927 has a mean sentence length of 16.8, very close to the COHA mixed material reference value, while “A Very Short Story” (1924) shows a sentence length of 15.6. Because of these scattered results it is unlikely that a pattern of shortening sentence structures from 1922 to 1928 can be sufficiently supported as there are results found showcasing both long and short sentence structures both in the first years of the period and the later. It is important to note that these extremes existed even at this early stage of Hemingway’s career. The mean sentence length for the entire period is 9.3, which is still far below the reference numbers, and even lower than the average for the entire Hemingway corpus. A text that is important to note during this period is *The Sun Also Rises* from 1926, which has a sentence length of 8.4.

Hotchner refers to the next period as The Key West Years, and it takes place between 1929 and 1936. After Paris, Hemingway would never again live in a large city, nor would he associate with a larger group of writers in the way he had done before. Instead, he began to surround himself with more common folk, who might not be as intellectually stimulating but helped him make his “simpler” characters become more alive (Hotchner 108). During the first year of this period *A Farewell to Arms* is published and provides us with a sentence length of 8.7, which is interestingly similar to *The Sun Also Rises* in this regard. The majority, close to 75%, of texts belonging to this period, were published in 1933 due to the publication of the short story compilation *Winner Take Nothing* in combination with a generally productive year for the author. Among these texts the highest frequency is found in “One Reader Writes” at 19.5. “One Reader Writes” is a concept piece, written in the form of a letter sent to a newspaper, which could explain why it deviates from the normative length of the period. It is also written by a female character, meaning that some of the oddities in comparison to Hemingway’s other texts might be due to the author attempting to simulate what he considered to be feminine language, or due to the story being only 3 paragraphs long. The



shortest mean structure length is found in “Homage to Switzerland”, also from 1933, at 6.7.

The results remain very broad, but both extremes of the period have moved towards a shorter norm, while the general mean sentence length for the period is slightly longer than during The Paris Years at 9.5 in comparison to the earlier 9.3.

The Spanish Years took place between 1937 and 1940, and continue the pattern of slight increases in general sentence length at 9.7. It also contains a novel, *For Whom the Bell Tolls*, which was released towards the end of the period in 1940 and has a mean sentence length of 9. The year of 1938 saw the publication of “The Denunciation”, 10.3, and “The Butterfly and the Tank” at 12.7 followed by “Night Before Battle” and “Under the Ridge” at 9.1 and 12.1 published in 1939. The novel of the period remains fairly close to the previously published novels, but the other texts seem to level out in terms of extremes. Even though the period in its entirety has provided slightly higher results than the earlier ones, the short stories published during the period are closer to both each other and the general period statistic in terms of sentence length than previously. This means that there are no extremes towards either long or short structures.

The Cuban and Venetian Years, Hemingway’s final periods represented in our corpus, shows some interesting breaks from the patterns that have been hinted at up to this point. The mean sentence length for the entire period is 11.3, which is higher than earlier periods and also higher than the results for the entire corpus. This is visible throughout all publications in the period with the shorter values being between 10 and 11, while “The Faithful Bull” (1951) has a longer mean sentence at 18. The novel of the period, *The Old Man and the Sea* from 1952, follows the pattern of longer sentences with a value of 13 words per structure.

The claims of shorter structures being a part of Hemingway’s style are correct throughout all the periods, as they are all showing values that are sufficiently different from the reference corpora for the trait to be considered as something that distinguishes

Hemingway's language from that of other writers. The CEN and COHA Fiction corpora suggest that there was a move towards shorter sentences in fictional literature from 1880 and forwards, but even in this regard Hemingway stands out as using shorter structures than the other writers.

There are some texts that provide results that deviate from this and in some cases even show lengths longer than the reference mean length, but these texts are the extremes of their individual periods. However, there does not seem to be any changes during the periods suggested by the critical material, while there is a noticeable change towards the reference mean sentence lengths during Hemingway's final period of activity (fig 5). If one considers a penchant for shorter sentences structures as a style feature related to Hemingway's journalistic roots, Dewberry's words about these being more visible in the author's early works is seen in the data above as well (25). This is especially true for the novels, perhaps indicating that they were written with a stricter adherence to the rule set than the short stories.

Type/token-ratio (henceforth TTR) would be the most suitable statistic to use for exploration into Hemingway's proposed obsession with counting words and repeating phrases as it shows how many times the types of a text is repeated amongst the tokens, or how many times a word is repeated within a text. This can then be standardized to account for different sizes of corpora, making comparisons possible. The standardized value is calculated by setting an N-value, being any number but normally 2,000, and then dividing the different corpora into sub-corpora of the same size. This accounts for the impact different sizes of text compilations can have on the statistic, as a larger text contains a higher number of tokens, by collecting the underlying data from several smaller samples. Finally, a TTR value is calculated for all the N-value sized sub-corpora separately and those values are combined into a standardized mean value for the entire corpus. Thankfully, this statistic is available in WordSmith's statistics tab, enabling us to get the data instantly. The Hemingway corpus

provides a standardized TTR of 35.42, which means that there is quite little lexical variation, or varied word choice, found throughout the texts in the corpus. The COHA Fiction corpus has a standardized TTR of 41.63 while the CEN has a ratio of 43.38. meaning that Hemingway's production has a lower variation compared to both. Through these numbers it is visible that Hemingway does, just as proposed, tend slightly towards a less varied and more repetitive vocabulary compared to other authors before him as well as his contemporaries. This means that a general assessment of Hemingway's authorship as being shaped by selective and repeated word choices would be correct, albeit to a smaller degree than one could have anticipated due to the pervasiveness of the description in the critical material.

Meyers noted that a change took place during 1934 where Hemingway moved further towards selective and conservative word choices, in combination with the previously explored shortening of sentence lengths. This will be approached by looking at the temporal sub-corpora, much in the same fashion as was previously done regarding sentence lengths. The data for each sub-corpus can then be compared to the overall type/token-ratio of the Hemingway corpus in order to establish where any change has occurred and what the change could be, as well as with the reference corpora in order to see the data in context.

During the Paris years, 1922 to 1928, Hemingway's output remains close to the general statistic with a standardized TTR of 34.8. The Key West years' sub-corpus, between 1929 and 1936, moves even closer to the number for the entire corpus at 35.2. There does not appear to be a sudden change directly after 1934, but the trend could be building up over time. The Spanish years, from 1937 to 1940, remain close to the previous period and the corpus general statistic at 35.1, rising slightly during the end of the author's life during the 40's and early 50's to a type/token-ratio of 37. The posthumous and unpublished texts remain close to the corpus statistic at 34.8. Much like with sentence length, there is a slight rise in language variation during the final publications, but it is still below the average variation given by CEN

and COHA Fiction. The changes between periods are quite small and do not show any specific periods of change, except for the arguable peak during Hemingway's final publications.

### Gender Participation in Hemingway's Writing

From the beginning of Hemingway's career, critics and authors such as Lillian Ross, Carlos Baker and Sheridan Baker have pointed out his often stoic depiction of male pursuits in different settings, and Hemingway's language is, by some, considered to provide an "understated masculine style" which permeates both his characters and writing (Sanderson 170). This style is perhaps influenced by "his selection of bullfighters, bridge blowers, lion hunters, and big-game fishermen", all professions and amusements heavily associated with a strong sense of masculinity (Weeks 5). The characterization of Hemingway's style by Weeks provides a solidly plausible definition of the male characters in Hemingway's fiction. This is important due to the narration almost exclusively taking place from one of these characters' perspective. Because of this it is necessary to note that the overly masculine nature of the characters could influence descriptive language from their perspective, for instance in regard to situations or other characters.

Starting in the 1930s and continuing onwards, Hemingway was so aware of the idea of the masculinity described in his writing that he himself constructed a persona as manliness incarnate based on the male ideals he himself had constructed. According to some critics, the female characters in his fiction suffered because of this, as they seldom held any individual personality (Meyers 427). The perhaps first instance of criticism towards Hemingway's way of treating women in his texts is found in Edmund Wilson's *The Wound and The Bow* from 1941, in which he argues that Hemingway tends to antagonize his female characters (235-237). The general contemporary consensus among the critics was that Hemingway deliberately focused on male characters that merely served as projections of his own

endeavors, only greatly exaggerated, as stated by critic Michael Reynolds, but others begged to differ (Sanderson 109).

Jeffrey Walsh's analysis of the narrative methods of *A Farewell to Arms* (1929), *For Whom the Bell Tolls* (1940) and *Across the River and Into the Trees* (1950) in *American War Literature: 1914 to Vietnam* (1982), surprisingly makes no mention of the way women are depicted. Instead, Walsh suggests that "the ordeals that Hemingway's male protagonists undergo can stand for universal human experience", and by suggesting this, Walsh goes against the discourse surrounding otherwise controversial female characters found in these three novels (Mazzeno 128).

The idea that Hemingway's style is being centered around the male view, meaning that Hemingway most often provides description from the perspective of a male character, is also touched on by Leslie Fiedler in Robert P. Weeks' *Hemingway: A Collection of Critical Essays* (1962): "Hemingway is only really comfortable dealing with 'men without women' " (86). In summary, the critical material makes a claim of Hemingway's use of male characters being dominant to a point of excluding female characters so much that it becomes a marker of style specific his writing. The most direct way to research this is to look at gender specific signifiers of character participation in the Hemingway corpus. The signifiers chosen to show participation in this project are "he" and "she" as they tend to collocate to verbs, which signifies participation rather than representation, and are gender specific.

"He", the most common signifier of definite masculine participation in a text, is found in the keyword list produced when comparing the Hemingway corpus to the Corpus of Historical American English mixed material reference corpus, but not when using the COHA Fiction as reference corpus. It occurs at a frequency of 9,906 in the Hemingway corpus which gives a standardized frequency of 1,191/100,000. To put this in context, the feminine equivalent signifier "she" is found at a standardized frequency of 264/100,000. "He" is also

highlighted in the 9th rank of the keyword list with a keyness of 1850.085 when compared to the COHA mixed material reference corpus. Looking at the density of occurrences within the texts it is visible that certain texts are far denser than others in this regard (fig. 6).



*Fig. 6: Concordance plot view of the masculine 3rd person pronoun "He" in Hemingway's novels.*

In figure 6 the instances of "He" appearing in the text has been plotted in a novel from every time period in the corpus. To the right side of the figure the number of hits, as well as text length, is also visible. This density of pronouns is not found to the same degree in Hemingway's short story production, suggesting that a different perspective, or a different type of narration, is at work in them. When comparing to the feminine 3rd person pronoun "She", Hemingway's preference for male characters is astoundingly visible, as the masculine variant is found close to, or higher than, the double frequency in each case<sup>2</sup>. There is, however, one exception to the rule of the general differences between novels and short stories.

---

<sup>2</sup>"Her" and "Him" were suggested to be included, but due to the relation between subject/object they are not as relevant to the concept of participation.



Fig. 7: Concordance plot view of the feminine 3rd person pronoun “She” in Hemingway’s novels.

The exception is “The Garden of Eden” (1987), which is considered an early version of *Islands in the Stream* (1970) but was edited heavily by Scribner’s editor Tom Jenks. The incomplete chapters were published some years later as a part of *The Complete Hemingway Short Stories* in 1987, and the story is included in its original, uncompleted form in the corpus used for this essay. The reason for the omission of the published version is due to the widely accepted notion that it is very clumsily and poorly edited, and does not reflect Hemingway’s style as much as that of the preferences of the editor (Donaldson 141).

While the frequencies of the masculine and feminine signifiers are far from similar in the novels and short stories published during Hemingway’s life, the posthumous “The Strange Country” contains 299 occurrences of “He” and 207 occurrences of “She” (Fig. 8). In some ways the findings could be considered indicative of a general change in the use of pronouns as a style marker towards the end of Hemingway’s active years. The density of occurrences also indicates a movement towards the use previously found in the novels. This section started with the masculine 3rd person singular due to its position among the keywords and the

significance of it as a marker of masculine activity in the text, but the finds indicate that in order to properly trace masculine action, this essay also has to take narrative style of the texts into further consideration.

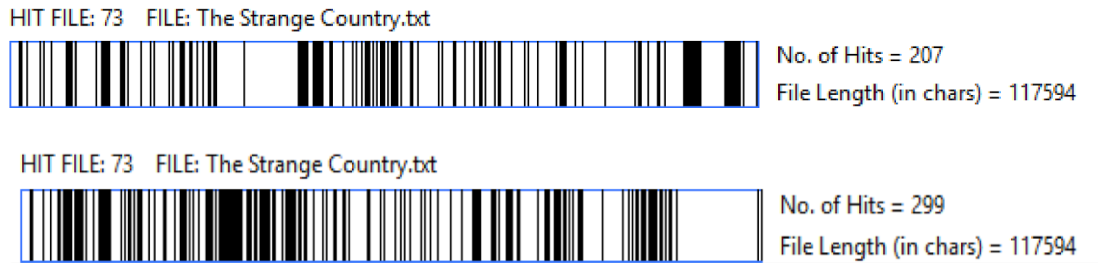


Fig. 8: Top: Feminine 3rd person singular pronouns in “The Strange country.

Bottom: Masculine 3rd person singular pronouns in “The Strange country.

Hemingway’s use of narrative as a tool for creating thematic effect has been written about earlier, and excellent examples of ideas concerning it can be found in Scott MacDonald’s “Narrative Perspective in the Short Stories of Ernest Hemingway” (1970) and E. M. Halliday’s “Hemingway’s Narrative Perspective” (1952), dealing with his short stories and novels respectively. Hemingway moved between first-person, third person utilizing character and third person omniscient throughout his different texts, and occasionally also within a single work (MacDonald 2). This indicates that the use of different narratives is a way for the author to create different effects, for instance the use of first person as immersive narration or third person as evaluative or descriptive, depending on the desired theme of the section.

The apparent style of fleeting narrative could be used for different ends, the change to first person present in *The Sun Also Rises* (1926) creates a “kind of moving now” that had previously been inconceivable for literary critics and Frederic Henry’s retrospective first person narrative creates an “aire of reminiscence” in *A Farewell to Arms* (1929), according to Halliday (MacDonald 3). MacDonald writes that “Hemingway not only experimented with the possibilities of narrative perspective, but experimented widely and successfully with



them” (256). A footnote in Carlos Baker’s *Hemingway: The Writer as Artist* claims that Hemingway did not start using third person narration until the 1930’s but MacDonald disagrees, and the concordance plots indicate that MacDonald is correct due to the pronouns related to third person narration being found in significant amounts in the material published before 1930 as well (4). Phillip Young’s *Hemingway: A Reconsideration* (1952) lists stories that make use of the first-person narrative, sometimes as a perspective provided by an observing character. However, Young also defines the narrative character when active in the plot as usually being a male protagonist similar to that of Nick Adams, who figures in 24 short stories written between 1920 and 1930 (57).

Returning to the previously presented results, even though “He” might have been the only gender specific pronoun highlighted in the keyword query, Hemingway’s use of different narrative perspectives means that any queries based solely on the 3rd person singular pronoun would omit a large amount of relevant data. Furthermore, Young’s conclusions regarding the nature of the characters hidden behind the “I” form allows for the 1st person nominative personal pronoun to be included in the list of masculine signifiers in the majority of the cases found in our corpus. The choice to include “I” in the list is also supported by the pronoun appearing on the fourth rank of the keyword list with a keyness rating of 5,557.286 and a standardized frequency of 1,607/100,000, making it a larger aspect of both the Hemingway corpus in general, and Hemingway’s style in comparison to the COHA, when compared to “He” and showing how much data would have to be omitted if it was removed as a signifier. This would, however, not be the case when discussing “The Strange Country”, as it was published after the works by both Baker and Young, and due to the shift in gender representation when looking at the 3rd person pronouns hinting towards a different gender participation frequency entirely.

As our analysis shows, Hemingway did indeed mainly deal with “men without women”, at least in the material published while he was alive. This is also to be considered a genuine marker of style, as both signifiers used for these observations are found on the keyword list when comparing Hemingway’s work to that of other writers active during his own century. An interesting note for further research on the topic would be “The Strange Country”, and *Islands in the Stream*, which do not show the same pattern. If a close reading was performed while keeping this data in mind it would be possible to present new insights and interpretations on Hemingway’s relation to the masculine and feminine towards the end of his career. Furthermore, the editing of “The Strange Country” and *Islands in the Stream* was, as previously mentioned, different from earlier works, so further clues as to why this change appears could perhaps be found there.

#### Gender Representation in Hemingway’s Writing

The critical discussion concerning the topic of whether or not Hemingway’s fiction can be considered misogynistic, sexist or overly negative towards women in his fiction has slowed down since the late 1980’s. Starting with Edmund Wilson in 1941, who did little more than note the “growing antagonism” perceptible towards the female, stating that “[l]ooking back, one can see at this point that the tendency has been there all along” (237), the debate became a great deal more heated with the surge of the women’s rights movement in the 1960s (Sanderson 171). Of course, by then there had been an increase in material from Hemingway, such as *Across the River and Into the Trees* (1950) and *The Old Man and the Sea* (1952) as well as a number of short stories and articles, and the negative trends could now be found more easily. The debates of the 60’s were, at the time, very unfavorable towards Hemingway and his reputation was diminished in academic circles. The attention the subject gathered had sparked the need for a rereading of Hemingway’s female characters, whose supposedly

inferior and one-sided position now came under intense scrutiny in contrast to the male character (Mazzeno 127).

Perhaps the most blatant of these controversial female characters, Catherine in *A Farewell to Arms*, would fall into the latter category of the dichotomy proposed in Roger Whitlow's *Cassandra's Daughters: The Women in Hemingway* (1984). According to his study, it had become common practice to place the women in his fiction within two brackets: the emasculating, dominant shrew or the submissive love-slave, which in turn was mirroring his own supposedly sexist mindset (Sanderson 171). That Catherine ultimately dies, delivering a stillborn son, was according to Judith Fetterley a "message to women readers" that "the only good woman is a dead one" (71), something which was challenged by Joyce Wexler in 1981. Wexler argued against the negative aspect of Catherine's character traits, demonstrating how she not only possessed a fully fleshed out personality to rival the male protagonist's (Frederic), but also became "Frederic's model of courage" (116), which in turn is corroborating with Sanderson's idea that "she is the creation of Frederic's bereaved memory" (Sanderson 182), being crucial to his idea of the perfect world.

Tying several of these ideas together, Mimi Gladstein's 1986 essay *The Indestructible Woman in Faulkner, Hemingway, and Steinbeck* provided a nuanced picture of the issue of Hemingway's proposed misogyny. Drawing on arguments produced by Elizabeth Janeway and Dolores Barracano Schmidt in 1971, as well as Leslie Fiedler in 1960, she agrees with Sanderson's claim regarding the submissive woman who is either a "bitch mother" or a "nurse mother" and that Hemingway projected his sexist mindset onto the female characters (Mazzeno 130). However, Gladstein does not see this as a necessarily malicious action born out of spite, but rather as a metaphor for a part of the world that cannot be controlled by the male desire to explore and conquer, a theme which is central to much of Hemingway's writing. As a result, the male character's inability to see a female as more than just an obstacle

on the road towards the ultimate goal of the hero is due to the hero's inherent inability to understand the symbolic value of the individualized female character (Mazzeno 131). The nature of Gladstein's work, which takes into account both the long debated notion of misogyny while also providing insight on possible explanations outside of the previously mentioned solutions focusing extensively on Hemingway himself is therefore highly valuable for this essay as it, too, looks at the text beyond the author.

The preferred method of investigating the representation of women and their actions in Hemingway's work with the help of corpus methods would be to pick out signifiers of female presence in the texts, similarly to what was previously done concerning pronouns when looking at participation. Furthermore, this section will also look at gender specific nouns, as descriptions of women from any perspective are relevant in order to define the representation of women in Hemingway's writing. The third person personal pronoun will be revisited, and the PoS-tagger will be used to look at different collocating verbs in order to see any recurring types of action performed by female characters. The nouns "woman" and "girl", in both singular and plural forms, will be explored in terms of collocating adjectives in the same fashion. This will show both the actions commonly associated with women in Hemingway's work and the ways that females are most commonly described by the author.

The PoS-tagged corpus will be searched for structures containing an article, an adjective and a female gendered noun. We attempted to find additional nouns to explore on this topic by looking at colloquialisms of the 1920's such as "gal", "broad" and "ankle", but found zero results in the corpus. We also looked for negatively connotated nouns but only found a recurring result for "bitch", which was exclusively found in the quadra-gram "son of a bitch" in all instances. "Woman" and "girl" will act as node words for a "skip-gram", which is an N-gram where one or more of the words have been replaced with a wildcard search for a specific part of speech, in this case an adjective (Pinna & Brett 110). This allows the query to

provide results on which type of adjectives are most commonly used to describe either noun. A similar skip-gram will be created for “she”, so that descriptions from different perspectives can be gathered.

The first queries are simple frequency statistics for “girl”, “woman”, “girls” and “women” to see how often females are described or interacted with, as the data for “she” mainly accounts for actions performed by female characters. “Woman” appears 252 times in the Hemingway corpus and is found in 30 out of the 69 texts it contains. Out of the texts *For Whom the Bell Tolls* accounts for most of the hits with 58 occurrences, a high number in comparison to the 1-3 hits found in the other texts. This could be explained by Hemingway’s new perspective on his relation to women during the period he was working on the text, in combination with a response to previous criticisms regarding the low number of female characters in his work (Sanderson 187). The plural form “women” is only found 83 times, but the results do not peak in frequency for any of the 27 texts where the word is found.

Finding a premodifying adjective is quite uncommon in the Hemingway corpus. An adjective being the L1, or first to the left, of the noun “woman” is found in 56 cases out of the 252 where the noun appears. The adjectives fall mainly into three different categories, presented according to frequency:

- 1: Adjectives dealing with age, such as “Old” or “Elderly” are the most common.
- 2: Positive attributes concerning looks or personality, such as “Fine”, “Good-looking”, “Honest” or “Admirable”.
- 3: Determiners of nationality, such as “American” or “Indian”.

There are also single uses of negatively connotated adjectives that would belong in the second category if they had appeared at a higher frequency, for instance “Fat”, “Wicked” and “Crazy”. There is also one use of a superlative adjective in this structure found in the corpus,

“the biggest woman”. There were no results for comparative adjectives in this structure. The results coincide with Harry Levin’s observation found in *Observations on the Style of Ernest Hemingway*, emphasizing “the restricted choice of adjectives and the heavy load of subjective implication carried by such uncertain syllables as ‘fine’ and ‘nice... it does not describe, it evaluates” (Weeks 77). When replacing the singular noun with the plural form, the query yielded 19 hits, where three had a frequency of more than one occurrence. The adjectives found in collocation with “women” fit into the same categories as those found for the singular noun, with the exception of adjectives indicating number, such as “many” or “several” which were both found on several occasions.

“Girl” is found at 589 instances in 42 of the texts, and thus seems to be Hemingway’s preferred female noun. 189 of these hits are found in *Across the River and into the Trees* from 1950. The two most common collocates in the L1 positions are the articles, with the definite article appearing 339 times and the indefinite 66. In the R1 position the highest frequency collocate is the verb “said” which was found 130 times, indicating that speech is the most common action performed by female characters. The verb with the next highest rating in collocation with “girl” is “asked” at 22 hits, which further implies a speech act being the most common action as it is the R1 collocate in 152 of the 245 instances the structure of “girl” followed by a verb appears. The types of adjectives found are very similar in nature to those found in collocation to “woman”, but the categories appear in a different order when sorted according to frequency:

- 1: Positive attributes concerning looks or personality, such as “Fine”, “Good-looking”, “Lovely” or “Grand”.
- 2: Determiners of nationality or ethnicity, such as “American”, “English” or “Jewish”.
- 3: Adjectives dealing with age, such as “Old” or “Elder”.

As with “woman”, there are lower frequency hits that describe negative attributes, such as “Crazy”. An interesting find was the use of “Forceful”, which did not appear in collocation to “woman”. This could be argued to hint towards agency, and the ability to act it implies, being something mainly connoting to youth, or even young women. “Forceful” only appears in the L1 position in the short story “The Butterfly and the Tank” from 1938, but is also used in *The Sun Also Rises* where it is again used to describe a female character from the perspective of Cohn as seen in the KWIC line below:

By that time, though, he had other things to worry about. He had been taken in hand by a lady who hoped to rise with the magazine. She was very forceful, and Cohn never had a chance of not being taken in hand. Also he was sure that he loved her.

It is also interesting to note that “elderly”, from the collocations of “women”, has changed to the comparative “elder” when being part of a structure containing “girl”. This makes good sense due to the word choice “girl”, implying youth in relation to the other feminine nouns, and thus needing the comparative form instead of the normal adjective. However, it seems to show a habit of word choice in relation to female characters having been changed in accordance with context rather than being replaced with a completely different adjective.

Moving on to looking at “She”, the starting point is a collocation query specified to find descriptive functions, in order to see if the descriptions of female characters from this perspective remains the same as they were in relation to the nouns. The results are similar when looking at adjectives found in position R2 of both the contraction “She’s”, which is Hemingway’s preferred use of “She is” in dialogue, and “She is”, which is the structure used in narration. The results indicate that the most common type of adjectives found here describe physical attributes, for instance “beautiful”, “attractive”, “lovely” and “too fat”. However, none of these were found at a significant frequency as the structure is exceedingly rare

throughout all periods. The same is visible in the same query with “She was”, which provides no results with more than one occurrence.

When looking at verbs in collocation to “She” a few interesting ones appear in the top 50. However, it is important to note that the vast majority deal with common actions such as “gave”, “walked” and “talked”. The ones that show more of the female behavior and the relation between male and female characters in Hemingway’s novels are, in order of frequency: “asked”, “smiled”, “loved”, “kissed”, “wanted” and “cried”. The first observation is that all the examples chosen, and a majority of the verbs found at higher frequencies, are in a past tense. This indicates that the “aire of reminiscence” suggested by Halliday could possibly stretch further than *The Sun Also Rises* and could be visible in more of Hemingway’s production.

The reason for “asked” being included in the examples has to do with the implications and connotations of the word, and due to it already having appeared at a high frequency in correlation to “girl”. It implies someone (in this case female due to the gendered pronoun) either asks for guidance, information or permission, and due to the predominantly male characters found in the Hemingway corpus, the further implication is one of female dependence on either the guidance, knowledge or permission of males in the text. This is supported by the KWIC-lines provided by AntConc, where the structure “She asked” is often either followed by a male noun variant or a first-person object pronoun which implies a male according to Young. “Asked” appears on rank eight of the collocates, at a frequency of 45 for the feminine third person pronoun. The masculine version of the structure appears at a slightly higher frequency of 65, which must be considered with the difference in frequency of female and masculine third person pronouns in mind.

The next three verbs, “smiled”, “loved” and “kissed”, are interesting due to their rankings in the top 50 list of frequencies. These verbs imply Hemingway’s female characters



really are quite loving in nature, or at least kind, or maybe that they are sexually available. “Smiled” might not be as explicitly connected to the idea of romantic interaction or love as “loved” or “kissed”, but it still provides a perspective on the representation of Hemingway’s female characters often being written as friendly, or inviting. “Loved” and “kissed” are more direct in this regard as they either imply an emotive connection between characters or show an action related to it. “Wanted”, on the other hand, shows that female characters are also written with desires of their own to a certain extent, showing that women in Hemingway’s text would, at least on a surface level, have some influence on the plots and situations in the texts, or the actions of the other characters. This influence, however, seems to take the role of desires and the ability to “ask” for them to fulfilled by someone else rather than any agency to fulfill them on female characters own accord.

Finally, “cried” appears towards the end of the list with a fairly low frequency. It shows a female reaction opposite to the one implied by “smiled”, and less common due to the low frequency ranking in comparison. However, it is one of the few negative emotive verbs found in connection to the female pronoun, and worthy of mention because of this as it shows that Hemingway’s writing is not completely one-sided in this regard. Although, one third of the occurrences of “cried” are found in “The Strange Country”, further hinting towards that story being especially interesting in terms of women in Hemingway’s writing.

For his data to be understood in context, we must consider what is found for the masculine versions of the node words “man” and “boy”. “Man” is found 1276 times in the Hemingway corpus and “boy” occurs 506 times. The author’s preference in terms of age connotation is the opposite when compared to the data on women/girls. Verbs commonly collocating with “man” are “said”, further indicating speech as a popular act in Hemingway’s writing, and “looked”. “Asked” is found, but only in 12 instances. The low frequency of “asked” in collocation to “man”, especially when taking the higher frequency of the node

word into account, makes it a very rare action for masculine characters when compared to female. Other items on the collocation list are “watched”, “told”, “thought”, “knew” and “called”, hinting towards depiction of the mental state or content of male characters being more common than it is with the female ones. “Boy” shares a similar list of collocating verbs, although “thought” and “knew” do not appear. Both “man” and “boy” appears at much lower frequencies in their plural forms.

In summary, the few times that women appear in Hemingway’s work they are written in a rather specific way. There are, as always, exceptions, but the exceptions show a pattern of appearing mainly in one specific text in most instances. Looking at the verb collocations with Whitlow’s dichotomy in mind shows some support of the hierarchy suggested by him, but less data that supports his “emasculating, dominant shrew” category. The submissive position of women suggested by “asked” in combination with the high rankings of both “loved” and “kissed” could be argued to provide some support for Whitlow’s “love-slave”, but due to it depending on a rather extreme interpretation of the importance of “asked” this study will refrain from attempting to support or dismiss that category of the dichotomy. However, the verb collocates of masculine signifiers tend to be of a more dominant nature, mainly considering “told” and “called”. It could also be argued that this dominance is further seen in the more frequent use of the male signifier “man”, connotating with adult age, while the preferred for females in the texts is “girl”, connotated to youth or even adolescence.

Wexler’s suggested role of women in relation to men, or Catherine in relation to Fredric, seems correct beyond *A Farewell to Arms*, as support for it is arguably visible in both the adjectives and verbs, the former mainly describing women as loving, kind and beautiful creatures in most instances and the latter not showing any frequent markers of aggressive or emasculating behaviour. Hemingway’s preference for male characters is seen here as well through male signifiers once again being found at higher frequencies. The preference for male

signifiers throughout makes a discussion on the function of the male gaze in relation to the adjectives found in collocation to female signifiers more poignant. While our method cannot address the subtleties of emasculating or dominating behavior in the texts with any certainty, it does show that the description of women should be viewed with the hierarchy suggested by the verbs in mind. A female character having to “ask” for any desired action to take place due to a lack of agency, while a male character is more commonly shown to collocate to “told” indicates that he not only has the agency to act in his own interest, but also the agency to direct the actions of others. Due to this, the adjectives used to describe women could be argued to show the necessity of women appearing friendly and appealing to the male actor in order for their plot to make any progress. This interpretation echoes not only De Beauvoir’s historical description of the women as a second sex, but it also shows how the gender performance described by Butler is facilitated by the patriarchal structure of society as described by Hemingway. Furthermore, it supports Fiedler’s claim of female characters mainly being written as dependent on a male.

### Discussion and Conclusion

The purpose of this essay has been to provide a summary of the state of academic discourse concerning the works of Ernest Hemingway in order to identify some of the most common and well supported claims pertaining to his style. After perusing written material by some of the most critically acclaimed scholars within the field, two claims stood out as most prominent. The first claim regarded Hemingway’s language patterns and his allegedly “hard boiled” writing style that favored simplicity and a concise prose, which in theory meant that the length of his sentences as well as use of descriptive language would be significantly different compared to other contemporary authors. Moving forward, the second claim revolves around the depiction of women and their attributes, as Hemingway attracted a large amount of attention on the subject where many claimed that he was prone to sexism.

After identifying and thoroughly analyzing these aspects, the appropriate method of exploration was selected. Instead of conducting a purely literary analysis by using only theory, a combination of corpus linguistic methods and literary analysis was applied on a framework of criticisms constructed around the most often cited scholars on Hemingway's authorship. The linguistic methods produced several types of quantitative data, which were then matched with the qualitative analysis found in the critical materials in order to provide a nuanced picture of the style and content, in terms of gender, in Hemingway's writing from a stylistics perspective.

Our analysis of Hemingway's style corroborates what has been described in the critical material in most instances. However, the author's style, in terms of language patterns, is much more nuanced than the general descriptions make it out to be. When considering noun and adjective use or sentence length, the data seems to indicate a certain conservatism in Hemingway's writing, as his results are closer to the COHA Fiction detailing 1900-1950, but always drift slightly towards the tendencies apparent in the earlier century shown through the CEN corpus. These features are specific enough when comparing Hemingway's writing to the reference corpora for them to be considered features of style.

The features of language variation, repetition and sentence length are seen to evolve towards closer resembling the average data for Hemingway's contemporaries during the finale of his productive decades, but still remain different enough in the terms of the attributes ascribed to them to be considered representative of Hemingway's "hard boiled" style. The data for sentence length as well as noun and adjective use seem to split into two categories evolving separately at different rates, albeit in the same manner. The short stories show a larger variation in almost all instances than the novels do, which is likely due to the greater amount of rewriting and editing that went into the process of writing the novels when compared to the shorter texts. The novels tend to be closer to the style as described in

Hemingway criticisms, which hints towards the novels being preferred for scholarly work, perhaps due to higher availability and commercial success, as the descriptions are most likely based on them rather than the short stories based on the data.

The queries on gender in Hemingway's writing produced some interesting results, both in support of and against descriptions found in the critical material. The male perspective was clearly shown to be the dominant one through gendered pronoun queries, especially when counting the genderless first person pronouns as a marker of male participation due to the likelihood of them referring to a male character. The exceptions to this rule were "The Strange Country" and *Islands in the Stream*, since they both contain a gender distribution that was far less one-sided. Reasons for this difference are beyond the scope of this study, but one could explore this as an aspect of the change in Hemingway's style towards his later publications which was evident in both language patterns and sentence lengths.

When exploring the activity and representation of female characters through skip-grams and collocates this study found two signifiers, "girl" and "woman", that served as node words for the further queries. The adjectives appearing in the skip-gram search showed three main categories of description found in Hemingway's work: Age, Positive physical or mental attributes, and Nationality. These were the main categories for both node words, but they appeared in different ranks of frequency and with minor changes such as, for instance, the comparative use of "elderly" while collocating "girl". The adjective "Forceful" appearing in collocation to "girl" is also interesting from this perspective, especially when it appears to often have a positive connotation. The low number of occurrences of "Forceful" in the Hemingway corpus makes any further interpretation of it as significant very speculative, but it could be used as a focal point for future close readings of the texts.

When looking at these results it is important to remember that the descriptions of women in Hemingway's texts are from a male point of view, and that the female characters are subjected to a male gaze both from the narrating characters and the author. This means that the quantitative data here cannot provide the same interpretation based on sub-text that a more interpretative method would be able to.

Moving on to the verbs found in structures containing the female signifiers, the results once again partially support and disagree with previous descriptions of Hemingway's work. To a certain degree the verbs are paralleled to the behavior types that could be expected with the adjective data in mind, but they also include some actions that either go against the picture provided by the adjective search, or elaborate on the female characters in relation to their male counterparts. The verb query provided some support for the categories of female characters proposed in the critical material, although the negative categories (the emasculating, dominant shrew or the mothering bitch nurse) lack support in the data. One could argue that the concordance line previously used to show the function of "Forceful" could also show what was intended by these descriptions as the dominant behavior of the female character and her considerations in regards to marriage, indicate a somewhat emasculating behavior by 1900's standards. However, no quantitative support for these tendencies were found in the corpus data.

Some of the verbs found in close proximity to the female node words suggest that a hierarchy is seen in the language of the corpus based on interpretation, regarding the implied conditions of the word choice, for instance "asked". This would likely be an indicator of female submission throughout Hemingway's work, which is a common claim in the critical material regardless of whether the author considers Hemingway's fiction to be misogynistic or not. The results for the gender related sections also highlights agency as an important factor when discussing women in Hemingway's literature. These findings, in combination with the

low amount of female characters seen participating in the texts through the pronoun data and the categories used to describe females, besides “Forceful”, support much of the criticism of gender roles in Hemingway’s writing, but no results or statistics were found to support the more extreme of these claims.

However, the lack of quantitative results in support of an interpretation does in no way invalidate that reading of the text. The greatest weakness of the method used for this project is the lack of qualitative evaluation, although some similar interpretations were made in relation to gender. The statistic results cannot replace a close reading performed by a well-read scholar, but should rather be used as a supplement to the interpretations created by readings supported by literary theory. Furthermore, they could also provide hints towards new entry points for analysis where the quantitative data stands out in any way.

In conclusion, Hemingway’s style is shown to be very nuanced, and shifted on different occasions throughout his career. However, many of the contextually supported periods given for specific changes were not reflected in the linguistic data. Instead, many features of Hemingway’s style were shown to have changed during his final decade of productivity. The case is the same for the participation, representation and description of females, which all remained largely steady throughout Hemingway’s production up to the later years. The concepts of description and representation were different from what was described in the critical material, but reasons for this can be argued to be both the amount of work conducted on the author’s novels when compared to the short stories, as well as the critics picking up on aspects of the texts that might not have been as visible in the purely language based material used for this essay. One must also consider the political and ideological nature of literary criticism, which is likely to have influenced the nature of what different scholars decided to describe, and the occasionally hyperbolic natures of their descriptions.

Based on the previously presented findings, it would be beneficial for further research and studies to look at the texts that do not conform to either descriptions or stylistic patterns when compared to Hemingway's entire production. For instance, the findings presented here indicate that a feminist comparative study of "The Strange Country" in relation to Hemingway's other texts or a narratological analysis of "The Garden of Eden" and *Islands in the Stream* focusing on narration and narrative characters are very likely to produce interesting results and new insights. Any research into the context of Hemingway's last decade in terms of influences, surroundings and events are also indicated to be of interest due to the prominent changes found during that period.



## Works Cited

- Anthony, Laurence. AntConc. 2014 (Version 3.4.4) [Computer Software]. Tokyo, Japan: Waseda University. Available from <http://www.laurenceanthony.net/>
- Anthony, L. 2015. TagAnt (Version 1.2.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from <http://www.laurenceanthony.net/>
- Baker, Carlos. *Hemingway: The Writer as Artist*. N.p.: Princeton UP, 1998. Print.
- Baker, Sheridan. *Ernest Hemingway: an Introduction and Interpretation*. N.p.: Holt, Rinehart and Winston, 1967. Print.
- Barry, Peter. *Beginning Theory*. 3rd ed., Manchester University Press, 2009.
- Barthes, Roland. "The Death of the Author", 1967 in Wood, Nigel & Lodge, David *Modern Criticism and Theory: A Reader*, 3d ed. London: Longman, 2008. print.
- Beauvoir, Simone de. *The Second Sex*. New York, Vintage, 1984.
- Butler, Judith. *Gender trouble: feminism and the subversion of identity*. Routledge, Taylor & Francis Group, 1999.
- Crawford, William J., and Csomay, Eniko. *Doing Corpus Linguistics*. Routledge, 2016.
- Culpeper, Jonathan, and Jane Demmen. "Keywords." *English Corpus Linguistics*, Cambridge University Press, 2015. pp. 90-106.
- Fanger, Donald. *Dostoevsky and Romantic Realism: a Study of Dostoevsky in Relation to Balzac, Dickens, and Gogol*. Harvard University Press, 1965.
- Friedan, Betty. *The Feminine Mystique*. W.W. Norton & Company, 1963.
- Greiner, Rae. *Sympathetic Realism in Nineteenth-Century British Fiction*. Johns Hopkins University Press, 2012.
- Hemingway, Ernest . "A Man's Creed." *Playboy Magazine*, 1963.
- Hemingway, Ernest. *A Moveable Feast*. New York, Charles Scribner's Sons, 1964
- Hemingway, Leicester. *My Brother, Ernest Hemingway*. N.p.: Pineapple, 1996. Print.

- Halliday, E M. "Hemingway's Narrative Perspective." *The Sewanee Review*, vol. 60, no. 2, 1952, pp. 202–218.
- Hotchner, A. E. *Hemingway and His World*. N.p.: Viking, 1990. Print.
- Jones, Christian, and Waller, Daniel. *Corpus Linguistics for Grammar: a Guide for Research*.  
Routledge, Taylor & Francis Group, 2015.
- Leech, Geoffrey Neil, and Mick Short. *Style in fiction: a linguistic introduction to English fictional prose*. Routledge Taylor & Francis Group, 2015.
- Lévy, Maurice. *Lovecraft: a Study in the Fantastic*. Detroit: Wayne State UP, 1988. Print
- Louw, W.E "Contextual Prosodic Theory: Bringing Semantic Prosodies to Life" in Heffer, C and Sauntson, H, ed. *Words in Context. In Honour of John Sinclair*, Birmingham, ELR, 2000.
- Lynn, Kenneth S. "Hemingway Hastily, Carelessly, Chaotically Edited : THE COMPLETE SHORT STORIES OF ERNEST HEMINGWAY ." *Los Angeles Times*, 29 Nov. 1987, [articles.latimes.com/1987-11-29/books/bk-25153\\_1\\_hemingway-scholars-publisher-s-note-finca-vigia-edition/2](https://www.latimes.com/1987-11-29/books/bk-25153_1_hemingway-scholars-publisher-s-note-finca-vigia-edition/2).
- MacDonald, Scott M. *Narrative Perspective in the Short Stories of Ernest Hemingway*. University of Florida, 1970.
- Mahlberg, M. (2013). *Corpus Stylistics and Dickens's Fiction*. London: Routledge.
- Mahlberg, M. "Literary Style and Literary Texts" in Biber, Douglas, et al. *Corpus Linguistics*:
- Mazzeno, Laurence W. *The Critics and Hemingway, 1924-2014*. Boydell & Brewer, 2015.
- Meckier, Jerome. *Hidden Rivalries in Victorian Fiction: Dickens, Realism, and Revaluation*. The University Press of Kentucky, 2015.
- Meyers, Jeffrey. *Hemingway: A Biography*. N.p.: Da Capo, 1999. Print.
- O'Halloran, K. (2007) "The subconscious in James Joyce's 'Eveline': a corpus stylistic analysis which chews on the 'Fish hook' ", *Language and Literature* 16 (3): 227-244.
- Pinna, Antonio and David Brett. Forthcoming. Constance and variability: Using PoS-grams to find phraseologies in the language of newspapers. In Kopaczyk, Joanna and Jukka Tyrkkö (eds.) *Applications of Pattern-driven Methods in Corpus Linguistics*. Amsterdam: John Benjamins. 115–130.

Scott, M., 2004, WordSmith Tools version 4, Oxford: Oxford University Press.

Sanderson, Rena . "Hemingway and Gender History." *The Cambridge Companion to Ernest Hemingway*, Cambridge University Press, 1996, pp. 170–196.

Sinclair, John. *Corpus, Concordance, Collocation*. 3rd ed., Oxford University Press, 1991.  
Investigating Language Structure and Use, Cambridge Univ. Press, 2007. P. 334 -361.

Spencer, Hannah (2011) Semantic Prosody in Literary Analysis: A Corpus based Stylistic Study of H. P. Lovecraft's stories. Masters thesis, University of Huddersfield. This version is available at <http://eprints.hud.ac.uk/12904/> (1/10 -17 21:00)

Walsh, Jeffrey. *American War Literature: 1914 to Vietnam*. 1st ed., Palgrave Macmillan, 1982.

Weeks, Robert P. *Hemingway: A Collection of Critical Essays*. Prentice-Hall, Inc., 1962.

Wexler, Joyce . "E. R. A. for Hemingway: A Feminist Defense of "A Farewell to Arms"." *The Georgia Review*, vol. 35, no. 1, 1981.

Whitlow, Roger . *Cassandra's Daughters: The Women in Hemingway*. Greenwood Press, 1984.

Wilson, Edmund . *The Wound and The Bow*. Riverside Press Cambridge, 1941.

Young, Phillip. *Hemingway: A Reconsideration* . G. Bell and Sons, 1952.

## APPENDIX A: Hemingway Corpus Content

<u>Name</u>	<u>Year</u>	<u>Type</u>	<u>Notes</u>
My Old Man	1922	Collection	Pub. 1923
Out of Season	1923	Collection	
Up in Michigan	1923	Collection	Edited 1938
A Very Short Story	1924	Collection	
Cross-Country Snow	1924	Magazine	
Indian Camp	1924	Magazine	
Big Two-Hearted River Part I	1925	Collection	
Big Two-Hearted River Part II	1925	Collection	
Cat in the Rain	1925	Collection	
Mr. and Mrs. Elliot	1925	Magazine	
Soldier's Home	1925	Collection	
The Battler	1925	Collection	
The Doctor and the Doctor's Wife	1925	Collection	
The End of Something	1925	Collection	
The Revolutionist	1925	Collection	
The Three-Day Blow	1925	Collection	
The Torrents of Spring	1926	Novel	Unavailable
The Sun Also Rises	1926	Novel	
A Canary for One	1927	Magazine	
A Pursuit Race	1927	Collection	
A Simple Enquiry	1927	Collection	
An Alpine Idyll	1927	Collection	
Banal Story	1927	Collection	
Che Ti Dice la Patria	1927	Collection	
Fifty Grand (1927)	1927	Collection	
Hills Like White Elephants (1927)	1927	Collection	
In Another Country	1927	Collection	
Now I Lay Me	1927	Collection	
Ten Indians	1927	Collection	
The Killers	1927	Collection	
The Undefeated	1927	Collection	
Today Is Friday	1927	Collection	
A Farewell to Arms	1929	Novel	
On the Quai at Smyrna	1930	Collection	Edited in 1938
A Clean, Well-Lighted Place	1933	Magazine	
A Day's Wait	1933	Collection	
A Natural History of the Dead	1933	Collection	
A Way You'll Never Be	1933	Collection	
After the Storm	1933	Collection	
Fathers and Sons	1933	Collection	
God Rest You Merry, Gentlemen	1933	Collection	
Homage to Switzerland	1933	Collection	
One Reader Writes	1933	Collection	
The Gambler, The Nun, and the Radio	1933	Collection	
The Light of the World	1933	Collection	
The Mother of a Queen	1933	Collection	
The Sea Change	1933	Collection	
Wine of Wyoming	1933	Collection	

One Trip Across	1934	Magazine	
The Tradesman's Return	1936	Magazine	
The Capital of the World	1936	Magazine	
The Short Happy Life of Francois Macomber	1936	Magazine	
The Snows of Kilimanjaro	1936	Magazine	
To Have and Have Not	1937	Novel	Unavailable
The Butterfly and the Tank	1938	Magazine	
The Denunciation	1938	Magazine	
Old Man at the Bridge	1938	Magazine	
Night Before Battle	1939	Magazine	
Nobody Ever Dies	1939	Magazine	
Under the Ridge	1939	Magazine	
For Whom the Bell Tolls	1940	Novel	
Across the River and into the Trees	1950	Novel	
The Faithful Bull	1951	Magazine	
The Good Lion	1951	Magazine	
The Old Man and the Sea	1952	Novel	
A Man of the World	1957	Magazine	
Get a Seeing-Eyed Dog	1957	Magazine	
The Last Good Country	1972	Magazine	Posthumous
A Train Trip	N/A		Unpublished
An African Story	N/A		Unpublished
Black Ass at the Cross Roads	N/A		Unpublished
Great News from the Mainland	N/A		Unpublished
I Guess Everything Reminds You of Something	N/A		Unpublished
Landscape with Figures	N/A		Unpublished
The Porter	N/A		Unpublished
The Strange Country	N/A		Unpublished