DOI: 10.53469/isshl 2024.07(02).19

# "Nolite te bastardes carborundorum": Corpus Stylistics and Language of Margaret Atwood's The Handmaid's Tale

## Fei-Yu Ren

Department of English, School of Foreign Studies, Zhongnan University of Economics and Law, Wuhan, Hubei

Abstract: This paper investigates the stylistic and linguistic characteristics of Margaret Atwood's The Handmaid's Tale, a novel of Canadian origin. Employing a corpus-based approach, the researcher aims to analyze its stylistic features in-depth, focusing on specific predetermined categories. Quantitative analysis was facilitated through WordSmith 8.0 software, enabling the identification of keywords in context, which serve as indicators of the linguistic style within the selected corpus. Results indicate that the novel employs simple language and concise sentence structures. Moreover, it demonstrates lexical diversity, encompassing various themes with a higher type-token ratio (TTR), indicative of a diverse vocabulary. Notably, there is a notable prevalence of first-person singular pronouns throughout the text. Regarding content word distribution, verbs are the most frequently utilized, followed by nouns, pronouns, adverbs, and adjectives—additionally, the analysis of keywords aids in uncovering stylistic and linguistic elements embedded within the text.

**Keywords:** The Handmaid's Tale; Corpus Stylistics; Margaret Atwood; Wordsmith Tools.

## 1. INTRODUCTION

This study mainly examines the linguistic and stylistic features of Margaret Atwood's dystopian novel The Handmaid's Tale (Atwood, 1985). Canadian author Margaret Atwood is hailed as the queen of Canadian literature, possessing a strong sense of female consciousness and ecological awareness (Yuan, 2020). Atwood's unconventional upbringing instilled in her a lasting reverence for the wilderness and education and a keen awareness of her identity as a Canadian (Jadwin, 2009). This is directly reflected in her writing, demonstrating a profound ecological consciousness and deep reflection on the survival issues of Canadian literature (Atwood, 1982). Her experiences in Afghanistan made her aware of the terror of Muslim religious fundamentalism, and she "noted the absence of women in the streets and markets" (Atwood, 2005).

One of her representative works, The Handmaid's Tale, has won numerous awards, including the Governor General's Award for Canadian Literature and the Los Angeles Times Book Prize. At the time of its publication, Publishers Weekly stated that "on the shelf of cautionary tales that have entered into modern folklore, it should have a place." The novel portrays a society characterized by religious fundamentalism, egocentrism, binary oppositional thinking among the social elite, overdevelopment of technology leading to irreparable environmental damage, rapid ecological deterioration, rampant infectious diseases, overuse of pesticides and herbicides, difficulty in clearing nuclear waste, a significant increase in the proportion of men and women losing fertility, and the prevalence of various diseases such as cancer and radiation sickness, resulting in severe ecological degradation. The degradation of the natural environment brings about the high-pressure policies of Gilead, dividing women into various ranks, forcing them to wear different colored uniforms, be responsible for other jobs, lead the same daily life, and undergo brainwashing regularly. Handmaids, as reproductive machines, undergo regular insemination ceremonies. For men, Gilead similarly divides them into categories, strictly regulates sexual behavior outside the social elite, controls pornography rigorously, strictly excludes homosexuality, and restricts sexual behavior strictly within the social elite, between men and women. Various violent crimes are rampant in society, leading to social unrest. Both men and women who violate the rules will be hanged at a particular time as a deterrent (Atwood, 1985).

The Handmaid's Tale should be read as a "critical dystopia" (Tolan, 2007: 156) and a "speculative dystopia" (Schneider, 2022:347). That is, it envisages a precarious future for humanity and/or the planet, stemming from the realities of our times (156), so critics should analyze it from a broader perspective to figure out the "intra-action" (Barad, 2007) of the textual world and the real political world. Scholars have also commended this novel for bringing attention to issues that feminists consider increasingly urgent (Dam & Polar, 2021; Atwood, 2020) to fight against a fresh misogynistic narrative that is currently circulating with seriousness and is embraced (and

employed) by politicians like "Donald Trump in the US, Bolsonaro in Brazil, and the far-right Spanish political party VOX" (Gámez & García, 2021:82). The value of The Handmaid's Tale today far surpasses that of a mere novel; it has become a symbol of contemporary feminism, even evolving into a phenomenon in its own right (Revelles-Benavente, 2021). It symbolizes activists' pursuit of gender equality, hopes of breaking binary oppositions, and demands for women's legitimate rights. Therefore, it's vital to delve deeply into this novel, figure out its unique styles, and provide solid suggestions for contemporary feminist movements.

This research adhered to the primary concepts McIntyre and Walker (2019) outlined in their work Corpus Stylistics Theory and Practice. Its analysis embraced the definition they proposed for Corpus Stylistics, which involves "the application of theories, models, frameworks from stylistics in corpus analysis" (pp.14–15). Employing techniques and tools from corpus linguistics, it merged them with stylistic models and theories. This collaboration provided a method to delineate, assess, and scrutinize the style of both literary and non-literary texts. To fulfill the objectives of this study, the data was structured into a corpus comprising the plain text of The Handmaid's Tale (referred to as The Handmaid's Tale Corpus). Subsequently, following Leech and Short's (2007) approach, the researcher conducted a stylistic analysis of this corpus. To the researcher's knowledge, there is a shortage of literature concerning the implementation of a corpus-based stylistic approach to analyze the language and stylistics of The Handmaid's Tale. However, scholars have explored Atwood's manipulation of narrative perspective as a stylistic tool (Savona, 2016; McNees, 2003).

# 2. METHOD

This research draws upon the insights proposed by Adil Jaafar & Abdul-Salam Jasim (2022) and Fischer-Starcke (2010) as foundational frameworks for conducting an extensive corpus-based examination of Margaret Atwood's dystopian novel The Handmaid's Tale. Employing a combination of quantitative and qualitative methodologies, this study scrutinizes the linguistic attributes of the dataset. Specifically, stylistic analysis is undertaken utilizing predetermined categories derived from the checklist delineated by Leech and Short (2007). As depicted in Figure 1, these categories encompass four primary dimensions, from which this study selects relevant sub-categories within the lexical and grammatical domains for analysis.

## 3. PROCEDURE

To achieve the objectives of the study, the researcher has performed the following steps:

- (1) Analyzing the length of sentences and words within each corpus.
- (2) Computing Type-Token Ratio (TTR) and Standardized Type-Token Ratio (STTR) to assess lexical diversity.
- (3) Determining the predominant part of speech utilized in each corpus.
- (4) Compiling word lists categorizing specific word classes within each corpus, including content words (nouns, verbs, adjectives, adverbs) and pronouns.
- (5) Employing selected categories from Leech and Short's (2007) checklist to analyze these word lists.
- (6) Conducting a keyword analysis to identify significant terms.

Leech and Short's checklist of style markers							
Lexical categories	General, nouns, adjectives, verbs, and adverbs						
Grammatical categories	Sentence types, sentence complexity, clause types, clause structure, noun phrases, verb phrases, other phrase types, word classes, and general						
Figures of speech	Grammatical and lexical schemes, phonological schemes, and tropes						
Context and	Context and cohesion						

Figure 1: Leech and Short's Checklist of Style Markers

## 3.1 Corpus Tools

To conduct the quantitative analysis, three software are involved in this study; the first is WordSmith 8.0 (Scott, 2021), which has a remarkable statistic that makes it different from other software. The second one is TagAnt 2.0.5 (Anthony, 2022) software, which annotates the corpus.

#### 3.2 Data

The present analysis employs a combination of quantitative and qualitative methodologies to scrutinize the designated corpora: The Handmaid's Tale Corpus (THTC), derived from the freely accessible online e-book version of Atwood's 1985 novel, comprising 96,976 tokens. This selection is justified by the novel's status as Atwood's seminal work, renowned as a cornerstone of feminist literature globally, and its significant impact on feminist movements.

The reference corpus utilized is the D section of Green's (2017) Corpus of the Canon of Western Literature (CCWL) (Version 1.0). This corpus operationalizes the notion of the Western literary canon based on Bloom's (1994) delineation of canonical works. Bloom's authority in literary scholarship, the extensive citation and influence of his work, and the explicit taxonomy of texts and authors provided in Appendix A of his book make it a compelling choice. The CCWL encompasses authors and works from ancient Greece to the 20th Century, encompassing English and non-English literature. The D section specifically includes all canonical works from the 20th Century, or "the Chaotic Age" (287), thus encapsulating contemporary literature akin to *The Handmaid's Tale*. The CCWL-D consists of 10,564,090 tokens. A comparison with this corpus enables the identification of the distinctive stylistic features of *The Handmaid's Tale*.

#### 3.3 Compilation of the Corpus

Given the corpus-based nature of this study, data collection proceeds through two primary stages. Firstly, compiling a corpus entails electronically copying and storing the E-book version of *The Handmaid's Tale* as a self-constructed corpus in "txt" file format (hereafter referred to as THTC). This facilitates ease of manipulation using corpus software and tools. Consequently, the characteristics of this corpus can be delineated as specialized (crafted for a specific research objective), written, monologic, and diachronic (spanning successive periods). Table 1 below outlines the attributes of the study corpus and the above reference corpus.

The subsequent step involves annotation, wherein each word in the corpus is tagged with its respective part of speech, a process known as POS tagging. To accomplish this, the researcher utilized the TagAnt 2.0.5 software, which is freely available for POS tagging.

**Table 1:** Corpus Information and Abbreviations

Abbreviation	Corpus	Size	Number of Tokens	Year
THTC	The Handmaid's Tale Corpus	1 Sample	96,990	1985
CCWL-D	The Corpus of the Canon of Western Literature-D Part	159 Samples	10,564,090	20th Century

# 3.4 Words and Sentences Length

Analyzing the lengths of words and sentences is an initial step in a stylistic analysis of the provided data to delineate its linguistic characteristics. To achieve this, WordSmith software is employed to quantify the words and sentences within the corpus. The WordList tool within the software's statistical section categorizes words based on their letter count. Sentences are enumerated to indicate the corpus's total sentence count, defined within the software as any sequence of words beginning with a capitalized word and ending with a period, exclamation point, or question mark, with sentence length determined by the word count per sentence.

According to Baddeley et al. (1975), short words typically comprise 1–5 letters, while long words contain six or more. Figure 2, depicted below, displays a screenshot captured from WordSmith 8.0 software, illustrating the categorizing words found in THTC.

In communication, whether written or spoken, words are fundamental components. Simplified and shorter words contribute to a less intricate and more accessible discourse, while longer and more complex words contribute to a

formal and intricate discourse. As shown in Figure 2, word lengths within THTC range from 1 to 16 letters, with 2, 3, 4, and 5-letter words being the most prevalent in THTC.

1-letter	2-letter	3-letter	4-letter	5-letter	6-letter	7-letter	8-letter	9-letter	10-letter	11-letter	12-letter	13-letter	14-letter	15-letter	16-letter
words	words	words	words	words	words	words									
5.011	17.377	20.680	20.551	12.123	7.607	6.012	3.347	2,252	1.091	528	224	119	46	18	

Figure 2: WordSmith Statistic of the length of the words in THTC

Based on Baddeley et al.'s (1975) classification, *The Handmaid's Tale* contains more simple and short words than long and complex words. The following facts can explain this result. Firstly, a novel is, according to Britannica, "an invented prose narrative of considerable length and a certain complexity that deals imaginatively with human experience, usually through a connected sequence of events involving a group of persons in a specific setting" (Burgess 2024). The objective of a novel is to articulate and portray either the author's envisioned reality or the tangible world. Using straightforward and concise language can enhance the readability and comprehensibility of a novel, thereby broadening its readership and augmenting its influence. Moreover, within this particular novel, Atwood incorporates numerous dialogues among characters. Baddeley et al. (1975) posited that shorter words are more readily memorized and recalled than longer ones, leading individuals to favor more straightforward vocabulary in verbal exchanges. The incorporation of easy and concise language in dialogue renders the novel more authentic and facilitates readers' comprehension of the primary traits of the characters. Figure 3 also confirms this result.

The average word length of THTC, as is shown in Table 2, is 4.14. This result is consistent with the above result, and both indicate that in this particular novel, Atwood tended to use simple and short words instead of long and complex ones. The average word length of THTC is slightly lower than its contemporary literary works(4.31), with a gap of 0.17. This may partly explain why this novel has had such a significant influence on global feminist movements, for example in America, when In June 2022, the United States Supreme Court officially overturned the Roe v. Wade case, which deprived legitimate abortion rights; lots of women marched through the city, yelling that the Gilead is coming.

The standard deviation of word length serves as a measure of the variability in vocabulary length. The SD of THTC (2.13) exhibits a slight decrease compared to CCWL-D (2.28), indicating a difference of 0.15. This finding suggests that the words utilized in *The Handmaid's Tale* demonstrate relatively uniform lengths.

Concerning sentence length, Sanyal (2006) surveyed several USA participants to get a precise idea of the readability of the sentence, what is regarded as an easy sentence to read, and what is considered a difficult one. The result

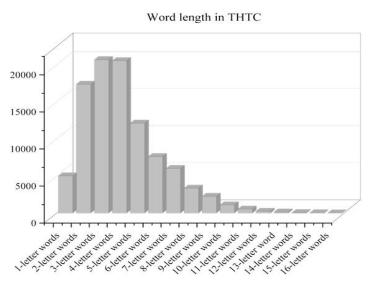


Figure 3: Word Length in THTC

**Table 2:** Average Word Length of THTC and CCWL-D

	Average Word Length	SD
THTC	4.14	2.12
CCWL-D	4.31	2.28

Such a study indicates that a sentence usually includes eight words is "straightforward" to read. In comparison, a sentence that consists of eleven words is considered an "easy" sentence to read, and a sentence that includes fourteen is considered "a fairly easy" sentence to read. Meanwhile, a seventeen-word sentence is regarded as a "standard" sentence. In comparison, the sentence with twenty-one words is considered "fairly difficult," and the sentence with twenty-four words is considered problematic. Lastly, a sentence with twenty-nine words is complicated (p. 63).

Regarding the length of the sentences in THTC, and by using WordSmith software, Table 4 is presented: Generally speaking, in contrast, a text with longer sentences is hard to understand, and often a reader with inadequate knowledge would misinterpret it, but longer sentences usually have a complex structure, rich details, and diverse contents, while readers easily and concisely understand a text with shorter sentences, but the meaning short sentences convey is not rich enough.

So, according to the information presented in Table 3 below, and depending on the classification that Sanyal (2006) has demonstrated in his study about the structure of sentences, the current researcher infers that the sentences in THTC are "fairly easy," which means that the language of handmaid's tale is relatively easy to understand. It includes easy and short words, but Atwood tends to use relatively longer sentences to convey more meanings to readers, leading to the fact that even though many scholars continuously interpret this novel from different aspects, more is needed to articulate its intricacies fully.

As for the standardized deviation of sentence length, the SD of THTC (9.77) is far lower than that of CCWL-D(21.70), with a gap of 11.93. Given that the reference corpus contains different genres of literary works, e.g., poems, prose, novels, sentence length, and structures vary distinctively among them. Therefore, this result indicates that Atwood used various sentence patterns in this novel, making it vivid and exciting.

Table 3: Average Sentence Length of THTC and CCWL-D

	Number of Sentences	Average Sentence Length	SD
THTC	2112	11.96	9.77
CCWL-D	631663	16.75	21.70

## 3.5 The Lexical Diversity

The Wordlist tool in WordSmith software can easily identify the type-token ratio of the words in each corpus. Token Type Ratio (TTR), according to Scott (2014), is "The ratio between types and tokens in the text" (p. 398). The TTR can be generated using WordSmith 8.0 by counting the number of tokens in the text (words) and the number of types in the text (only the different word types are counted, and all the repeated ones should be deleted). It can be considered a valuable analysis for an equal-sized corpus only because it highly depends on the text size, and the result of such analysis between two corpora that vary in length would be inconsistent. Besides, in practice, the vocabulary used in a language is limited. When the number of forms grows to a certain extent, the number of types may not continue to increase proportionally and may even cease to increase. Therefore, by segmenting texts into specific lengths (this paper uses 1000 as the calculation basis), the calculated Standardized Type/Token Ratio (STTR) can better reflect the richness of vocabulary used in longer texts.

The higher the Type-Token Ratio (TTR), the more diverse and extensive the vocabulary usage; conversely, the lower the TTR, the more monotonous and limited the vocabulary usage. As is shown in Table 4, the STTR of THTC is 43.80, illustrating that the words used in this novel are diverse and various. Compared with CCWL-D, this novel is more lexically diverse, with a gap of 0.46%. This finding indicates that Atwood is fairly talented and experienced in writing, and her educational background equipped her with rich knowledge and various sources (Yuan, 2020).

Table 4: Type-token Ratio and Standardized Type-token Ration of THTC and CCWL-D

Table 4. Type token Ratio and Standardized Type token Ration of Tittle and Cewil B									
Number of Types	TTR (%)	STTR (%)							

THTC	9037	9.32	43.80
CCWL-D	97418	0.92	43.34

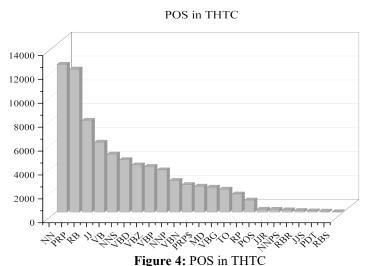
#### 3.6 Lexical Distribution

According to McIntyre and Walker (2019), merely assessing the lengths of words and sentences within a specific corpus may not sufficiently encapsulate its content; instead, a more practical approach involves discerning the various types of words present and identifying the most frequently occurring types, which aids in the identification of stylistic nuances within a text. Additionally, examining lexical diversity and focusing on the distribution of content words can shed light on whether the text exhibits a nominal or verbal style (formal or informal), a significant aspect of textual style amenable to stylistic analysis.

In the context of the present dataset, the TagAnt software is utilized to annotate each word in the corpus with its respective part of speech. The most straightforward method to explore the different parts of speech within a corpus is to generate a chart displaying the types of words encountered and their respective frequencies within the corpus under examination. Figure 4 depicts the word types identified in THTC and their corresponding frequencies.

After presenting a clear image of the linguistic diversity in the language of *The Handmaid's Tale*, it is essential to shed light on the distribution of content words<sup>1</sup>, central to conveying lexical meaning, including nouns, content verbs, adjectives, and adverbs, in particular, because that will help in characterizing the text into a nominalize or verbalize text style to indicate whether *The Handmaid's Tale* are nominal preferences or verbal preferences.

#### 3.7 Distribution of content words



**Table 5:** Content Words Distribution of THTC

	Noun		Verb		Adjective		Adverb		Pronoun	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%
THTC	19616	20.24	20405	21.03	5856	6.04	7853	8.10	14138	14.58

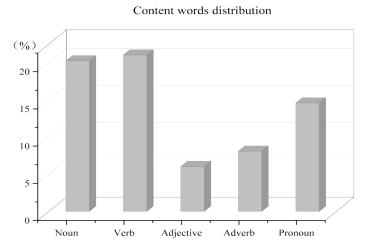


Figure 5: Content Words Distribution in THTC

Utilizing WordSmith software, the annotated data, i.e., data processed by TagAnt software to assign each word its respective part of speech, has been organized into distinct Wordlists. Each Wordlist comprises entries corresponding to specific word classes along with their frequencies of occurrence within the research corpus. The distribution of content words is determined by computing the frequency of occurrence for each type of content word.

As is shown in Table 5, in THTC, verbs emerge as the most frequently encountered category among content words (20405 Freq.). Upon sorting the verbs list by item frequency, verbs such as "be," "have," "do," and "say" demonstrate the highest frequencies. Nouns are the second most prevalent type of content word (19616 Freq.). Upon sorting the verb list by item frequency, nouns such as "time," "way," "woman," "aunt," "room," and "things" exhibit the highest frequencies. Additionally, pronouns (14138 Freq.) occupy the third position, with words like "I," "it," "she," "her," "they," and "he," displaying the highest frequencies in the pronoun list. Adverbs (7853 Freq.) constitute the fourth category, with words like "not," "now," "too," "so," and "then" exhibiting the highest frequencies in the adverb list. Finally, adjectives (5856 Freq.) rank last, with words like "other," "little," "white," "red," and "own," showcasing the highest frequencies in the adjective list.

From the above Figure 5, verbs exhibit the highest frequency of usage in THTC, followed by nouns and pronouns with relatively higher frequencies. In comparison, adverbs and adjectives demonstrate lower frequencies in this corpus. Thus, given the prominence of verbs compared to nouns and pronouns, it is evident that the language of *The Handmaid's Tale* is of verbal preference.

A preference for verb language suggests a relatively vivid and simple tone in Atwood's writing. Additionally, nouns rank as the second most prevalent category among content words. Integrating the preceding discourse suggests lively dialogues within the novel, characterized by a formal yet not overly rigid language but rather dynamic and engaging. Pronouns occupy the third position, implying a profusion of vivid and concrete character descriptions. The comparatively infrequent utilization of adverbs and adjectives suggests Atwood's inclination towards straightforward language, consistent with the memoir-style narrative, enhancing the novel's realism.

# 3.8 The detailed analysis of the wordlists

For further information about each type of content word used in this corpus, the researchers of this current study have analyzed the Wordlists based on the selected categories chosen from Leech and Short's (2007) checklist. The results are as follows.

The Verbs: Within this novel, various transitive and intransitive verbs are employed in varying proportions, which are crucial in conveying sentence meaning. Nonfactive verbs are more prevalent than active ones, with verbs such as "be," "feel," "think," "remember," "want," "look," and "wish" frequently utilized.

The nouns: Atwood tends to use concrete nouns, referring to concrete things and people in the Gilead, more frequently than abstract nouns; nouns like "time," "way," "women," and "aunt" are commonly used. Time is an essential term in this novel. As a "critical dystopia" (Tolan, 2007: 156) and a "speculative dystopia" (Schneider, 2022:347), that is, it envisages a precarious future for humanity and/or the planet, stemming from the realities of our times (156), this novel firstly shaped a pre-Gilead society, which is developed and free, and then due to the environmental pollution, Gilead came into power. The pre-Gilead period is the present of some countries and the future for others, and the Gilead period is the future of the future. The protagonist, Offred, often recalls the past pre-Gilead period, constructing a bridge between past, present, and future.

The Pronouns: Pronouns rank third in frequency of use, suggesting an abundance of vivid and detailed character descriptions. Written from a first-person perspective, the novel serves as the memoir and monologue of the protagonist, Offred, with the pronoun "I" being the most frequently employed. The utilization of "I" immerses readers into Offred's internal world, allowing them to share her struggles, resistance, and awakening. This narrative perspective fosters reader engagement with the protagonist, enabling them to empathize with her experiences.

The Adverbs: Analysis reveals that the novel incorporates various adverbs in varying proportions. Adverbs provide additional information about sentence elements, including verbs, adjectives, other adverbs, sentences, or paragraphs. The most commonly used type is additive/restrictive adverbs, followed by adverbs of time, linking adverbs, degree, manner, and place adverbs, respectively.

The Adjectives: Given the novel's utilization of simple and direct language, the author employs adjectives to enhance the impact of her writing on readers. Non-gradable adjectives are favored over gradable ones, and attributive adjectives are preferred over predictive ones.

# 3.9 The Keyword Analysis

Keywords, as elucidated by Scott (1997), denote words that exhibit a statistically significant increase in frequency within a particular text or corpus compared to a larger, comparable reference text or corpus. Their heightened occurrence is attributed to their significance in the analyzed data's content or structure, indicating prevalent topics or structural elements and thereby elucidating the data's meanings. Consequently, scrutinizing keyword usage can be an initial step in the literary interpretation of data.

The discrepancy in size between the two datasets forms the basis for inferring that the words identified as keywords hold importance for the content or structure of the data. The larger scale of the reference corpus diminishes the likelihood that the frequency of keywords in the data stems from general language usage probabilities, instead indicating that their presence is motivated by the data's structure or content. Consequently, hypotheses regarding the content and structure of the data can be formulated based on the keywords. Subsequent analyses of the keywords' concordance lines provide evidence either supporting or refuting these hypotheses.

Concerning this study, the target/study corpus used is the THTC, which includes 96976 tokens. As for the reference corpus used in this study, the researcher has selected an already existing dataset presented by Green (2017), namely the D section of CCWL. This corpus includes canonical literature from the 20th Century in 10 languages (non-English works are translated into English).

A Keywords program, such as the one integrated within WordSmith Tools, typically allows users to configure multiple parameters. Among these parameters is a minimum frequency cutoff point to exclude words deemed unusual solely due to their infrequent occurrence or non-occurrence in the dataset or reference corpus. Another parameter involves a test for statistical significance, which evaluates the importance of the unusualness of the identified keywords. As noted by Scott (2008: 145-6), in cases where selectivity outweighs the concern for risk, users might opt for a relatively low p-value threshold, such as 0.000001 (equivalent to 1 in a million or 1E-6 in scientific notation), to yield a smaller set of keywords. Criteria this research set for analyzing keywords are in line with that of Culpeper's (2009), including (1) achieving a quantity that aligns with the researcher's objectives, (2) avoiding an excessive volume of words for analysis, (3) ensuring a reasonable distribution of instances of even the least common keywords, and (4) minimizing the inclusion of isolated or scarce word types.

The researcher must configure the statistical parameters within the software to facilitate the analysis of keywords and generate a keyword list. In this study, the statistical settings employed were as follows: utilizing approximate

Bayes factors to gauge effect size, set at 33.00 degrees, with an associated p-value of 0.0000000001 and LL 40.66, aimed at providing robust evidence against the null hypothesis, which posits that any discrepancy in the frequency of a specific item between the two corpora arises due to chance. Employing the log ratio as a criterion for refining items within the keyword lists, any item with a log ratio equal to or greater than 1.5 is deemed significant. A qualitative analysis of its constituents should be conducted after configuring the necessary statistical tests for generating the keyword list. Figure 6 shows a screenshot of the SNC keywords list used in this study.

The word "we" is the key in the keywords list. From the concordance lines, as is shown in Figure 7 below, the word "we" stands for the handmaids. In the Gilead society, women are classified into different groups based on their functions. Handmaids act as surrogate mothers to parent sterile commanders' couples. The emphasis of "we" is dissipating of "I," that is the loss of individual identity. Handmaids wear the same red clothes, say the exact words, do the same things, and have the same destiny. In the hyperreal patriarchy-as-terror-regime (Dam &Polak, 2021:172), women are objectified and lose their own identity. Their fates are in line with that of nature in an anthropocentric society: negative, passive, oppressed, exploited, and persecuted.

Culpeper (2009) categorized keywords into three distinct groups. Firstly, proper nouns were identified, with Scott arguing their limited significance by illustrating how a text concerning racing might mistakenly highlight incidental mentions of horse names as significant (Scott, 2008: 143). Secondly, keywords associated with the text's thematic focus, or "aboutness," were delineated, a concept linked by Scott to Halliday's (1994) ideational metafunction, suggesting they are words recognizable or predictable to human readers (Scott, 2008: 143; Scott, 2000: 160). Lastly, indicators primarily reflective of stylistic features rather than thematic content were highlighted, with Scott providing examples such as "because," "shall," and "already" (Scott, 2008: 143). The style appears to encompass terms unrelated to thematic content (Scott & Tribble 2006: 60).

N	Key word	Freq.	%	Texts	RC. Freq.	Rc. %	BIC	Log_L	Log_R	D-Diff	P
1	WE	673	0.7%	1	24,815	0.23%	543.53	559.71	1.56	0.17	0.000000000
2	SAY	325	0.3%	1	11,039	0.10%	289.47	305.65	1.68	0.19	0.000000000
3	SAYS	308	0.3%	1	2,674	0.03%	946.23	962.41	3.65	0.12	0.000000000
4	US	275	0.3%	-1	9,497	0.09%	236.30	252.49	1.66	0.18	0.000000000
5	IT'S	226	0.2%	1	8,660	0.08%	160.58	176.77	1.51	0.15	0.000000000
6	IT'S	185	0.2%	1	0		1,722.72	1,738.91	116.97	0.24	0.000000000
7	AUNT	162	0.2%	1	1,393	0.01%	492.72	508.91	3.66	0.13	0.000000000
8	WOMEN	147	0.2%	1	3,375	0.03%	205.61	221.79	2.25	0.21	0.000000000
9	FEEL	137	0.1%	1	3,472	0.03%	169.97	186.15	2.10	0.17	0.000000000
10	AROUND	135	0.1%	1	1,696	0.02%	319.97	336.16	3.12	0.13	0.000000000
11	USED	130	0.1%	1	3,488	0.03%	149.44	165.62	2.02	0.19	0.000000000
12	MOIRA	125	0.1%	1	18		1,050.84	1,067.02	9.56	0.43	0.000000000
13	DIDN'T	118	0.1%	1	3,563	0.03%	114.01	130.19	1.85	0.21	0.000000000
14	RED	117	0.1%	1	3,462	0.03%	116.27	132.45	1.88	0.22	0.000000000
15	CAN'T	109	0.1%	1	4,196	0.04%	68.43	84.62	1.50	0.19	0.000000000
16	LYDIA	99	0.1%	1	41		745.80	761.99	8.04	0.11	0.000000000
17	THERE'S	93	0.1%	1	2,936	0.03%	80.61	96.79	1.79	0.17	0.000000000
18	FRONT	91	0.1%	1	2,200	0.02%	113.94	130.12	2.17	0.23	0.000000000
19	REMEMBER	89	0.1%	1	2,570	0.02%	87.68	103.86	1.92	0.22	0.000000000
20	SHE'S	88	0.1%	1	1,744	0.02%	136.84	153.02	2.46	0.12	0.000000000
21	JANINE	84	0.1%	1	0		773.38	789.56	115.83	0.27	0.000000000
22	LUKE	83	0.1%	1	47		594.72	610.90	7.59	0.42	0.000000000
23	THERE'S	73	0.1%	1	0		669.98	686.16	115.62	0.22	0.000000000
24	HARD	73	0.1%	1	2,713	0.03%	43.78	59.96	1.55	0.22	0.000000000
25	TURN	69	0.1%	1	2,530	0.02%	41.70	57.89	1.57	0.21	0.000000000
26	ISN'T	66	0.1%	1	2,231	0.02%	46.32	62.50	1.69	0.16	0.000000000
27	SHE'S	63	0.1%	1	0		575.99	592.17	115.41	0.22	0.000000000
28	SERENA	63	0.1%	1	28		464.16	480.34	7.94	0.63	0.00000000
29	DOESN'T	61	0.1%	1	1,520	0.01%	68.25	84.43	2.13	0.16	0.00000000
30	SOMETIMES	61	0.1%	1	2,287	0.02%	33.23	49.42	1.54	0.28	0.00000000
31	WALK	61	0.1%	1.	1,924	0.02%	47.38	63.57	1.79	0.24	0.00000000
32	WASN'T	59	0.1%	-1	1,246	0.01%	80.49	96.68	2.37	0.21	0.00000000
33	I'M	59	0.1%	1	0		538.39	554.57	115.32	0.24	0.00000000

Figure 6: The Keyword List of THTC Corpus

For the first category, there are six words in the keywords list, shown in Figure 6, i.e., "Aunt," "Moira," "Lydia," "Janine," "Luke," and "Serena." They are names of characters in the novel, and their positive keyness indicates their frequent appearance in the book and their significant roles in the story's development. However, as Scott(2008) said, their significance is limited.

Regarding the second category, eight words in the keywords list are related to the text's "aboutness," i.e., "women," "feel," "around," "used," "red," "remember," "hard," "turn," and "walk."

The word "women" indicates this novel's theme: women's survival in patriarchal hegemony. Male chauvinism and heterosexual hegemony objectify and victimize the female body, leading to an identity crisis or the loss of identity of female characters and exploitation of female bodies and psychological pressure. Under the pressure of fundamentalist religion, Aunts, and male society, fertile handmaids have to be devout to their duties as surrogate mothers to parent sterile commander couples to suffer the consequences of uncontrolled sabotage on the environment. The issue of how women survive in patriarchal societies has been a constant concern for Atwood, who also metaphorically maps this relationship between oppression and the oppressed onto the issue of survival of the Canadian nation, as a former British colony, and Canadian literature under the shade of American literature, thus shaping her strong sense of survival (Atwood, 2022).

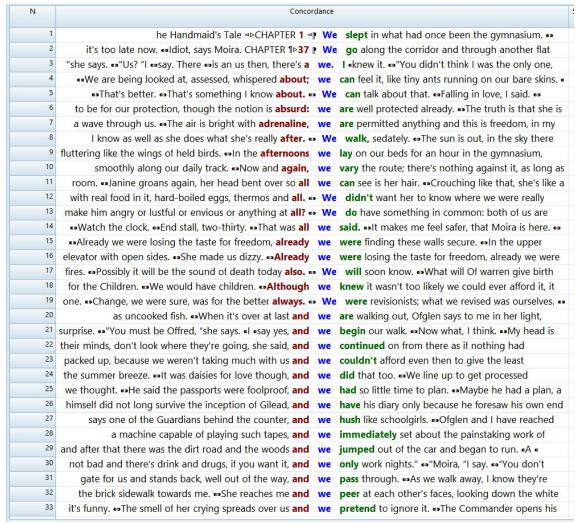


Figure 7: The Concordance Lines of Pronoun "we"

The terms "feel," "around," "say," and "used" also serve as poignant indicators of women's circumstances within the dystopian Gilead society, as discerned from the concordance lines depicted in Figure 8. Offred is compelled to acclimate herself to the stringent regulations imposed by the Gilead regime, reduced to a mere vessel, akin to a "used oven," tasked with the singular purpose of procreation. In conjunction with the term "memory," "used" additionally encapsulates Offred's recollections of her past life within the pre-Gilead society, engendering a stark dichotomy between liberty and subjugation and delineating the disparate societal statuses of women.

The color "red" has a symbolic meaning in this novel, as Atwood put it:

I... knew that clothing is a symbol, that all symbols are ambiguous, and that this one might signify a fear of women or a desire to protect them from the gaze of strangers. But it could also mean more negative things, just as red can mean love, blood, life, royalty, good luck—or sin (Atwood, 2005: 206).

Red, symbolically evoking connotations of vitality, fertility, and sensuality, emerges as a multifaceted emblem within societal constructs. In the Gilead regime, red assumes a distinctive significance as it becomes synonymous with the role and dentity of Handmaids. Within the confines of Gileadean indoctrination, women, often those who are biologically capable of motherhood, undergo a rigorous process of persuasion and instruction at the Red Center, wherein they are groomed to assume the role of Handmaids before their deployment into the households of the "global majority of males" (70). The educational framework prevalent within Gileadean institutions reinforces the prevailing narrative of male superiority, thereby perpetuating the systemic subjugation of women. Furthermore, the prominence of red also serves a pragmatic function, making individuals adorned in such hues conspicuous, particularly in contexts of flight or escape.

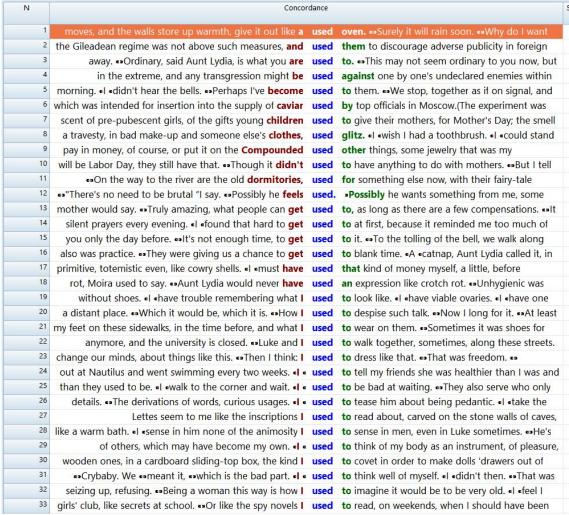


Figure 8: The Concordance Lines of "used"

Regarding the third category, words like "it's," "don't," "can't," "there's," "front," "she's," "isn't," and "I'm" belong to this category. Despite being a dummy subject, "it" denotes non-human beings or "not-me." This represents Atwood's consciousness of nature and women. "Only when the self directs its attention back to the object, regarding the self as an object, does the subject destroy itself?" The destruction of the subject's self is initiated by objectification. In the novel, the objectification of animals similarly foreshadows the subject's objectification of the self, marking the destruction of the subject's self-subjectivity. Offred and Luke, as sacrificial victims of power, are compelled to objectify and otherize their pet cat to facilitate their escape, ultimately leading to its demise. This symbolizes the brainwashing and oppression exerted by the Gileadean regime upon ordinary individuals. Under

such a system, people inflict harm upon each other, women are pitted against each other, and the relationship between humans and animals deteriorates markedly.

In this category, all the auxiliary verbs, model verbs, and linking verbs all appear in negative forms, such as "don't," "can't," and "isn't." They indicate that the Gileadian regime strictly prohibited behaviors and speeches that didn't agree with the fundamentalist interpretation of the Bible and its totalitarian policies. Besides, the word "there's" depicts the environment where handmaids live, the things they do, and the people they meet, which also exhibit the problematic circumstances under white terror.

The word "I'm," opposite to "we," illustrates Offred's realization and refind of her identity. Awareness of one's own body is the first step in the awakening of gender consciousness and the development of self-awareness. Offred recognizes her body as her territory, a domain entirely governed by feminine rules and culture rather than patriarchal norms, characterized as "nimble, singular, strong, and faithful to me." The body represents the final frontier for women under the oppression and assault of patriarchal culture, serving as the absolute other to men, something that cannot be indeed known or understood, remaining mysterious and steadfast (Atwood, 1985).

#### 3.10 Discussion

The research reveals that *The Handmaid's Tale* employs simple and accessible language to enhance the novel's readability and comprehensibility, thereby expanding its readership and impact and imbuing the narrative with authenticity, aiding readers in grasping the fundamental traits of the characters. The relatively low standard deviation indicates a consistent word length throughout the text. Atwood employs straightforward sentence structures characterized by brevity. Despite using simple and concise vocabulary, Atwood constructs longer sentences to convey nuanced meanings, necessitating a deeper exploration of the novel's complexities. A higher standard deviation suggests a diverse array of sentence structures, contributing to the vividness and intrigue of the narrative. The relatively high type-token ratio (STTR) in *The Handmaid's Tale* indicates lexical diversity, underscoring Atwood's adeptness in writing, bolstered by her extensive knowledge and diverse sources.

Regarding content word distribution, verbs are the most prevalent, followed by nouns and pronouns, with adjectives and adverbs occurring less frequently. This preference for verb-centric language imbues Atwood's writing with a lively and straightforward tone, characterized by vibrant dialogues, formal yet dynamic language, vivid character portrayals, and a narrative style akin to memoirs, heightening the novel's realism.

Keywords within the novel are classified into three categories: proper nouns, thematic keywords, and stylistic indicators. Through analysis, the narrative initially delineates the hegemonic rule of Gilead and the plight of women under its oppressive regime, followed by Offred's journey of self-discovery and her defiance against male dominance.

Despite her concern about the survival of women, Atwood also conveyed concern about the survival of all human beings. In a society marked by gender inequality, both sexes suffer. Women endure societal oppression and exploitation, while men, in constructing their societal roles as the dominant gender, exclude any elements that could be considered weak or insufficient, such as mental health issues. Human exploitation and oppression of nature inevitably provoke natural resistance, bringing about disasters that humanity struggles to confront. Atwood's concern for nature and others deserves significant attention, as each catastrophe has the power to alter our way of life, instilling fear and anxiety and revealing humanity's inherent flaws, exposing the atrocities of patriarchy and anthropocentrism time and again.

Butler (2020) perceives humanity's existence in society, recognizing the precariousness of human life and the intimacy and interdependence among individuals, and believes that the obligation to nonviolence lies not in romantic love but in the fundamental reliance on one another. General Secretary Xi Jinping proposed establishing a "Community with a Shared Future for Mankind," viewing human society as an interdependent community where nations rely on each other. Since humanity resides in a "global village," citizens of all countries are also citizens of the world, and global interests align with individual interests. Embracing sustainable development has transitioned from an idea to the Chinese government's actionable agendas and concrete plans. Upholding a global governance perspective and solving global issues entails a collaborative process involving governments, intergovernmental organizations, NGOs, multinational corporations, etc (Qu, 2021). Strengthening international norms and mechanisms is crucial in forming a mechanism with binding and ethical standards capable of addressing global issues, offering a uniquely Chinese approach to global governance. Beyond interactions among international

organizations, fostering a "community of shared destiny" between humans and nature is essential for harmonious and sustainable development.

# 4. CONCLUSION

This research aims to conduct a corpus-based stylistic examination of *The Handmaid's Tale*, written by Canadian author Margaret Atwood, to scrutinize its linguistic and stylistic characteristics. The analysis has been executed using a predefined checklist derived from Leech and Short (2007), along with corpus tools and software. The findings underscore the utility of corpus tools in facilitating stylistic analysis, streamlining the process, and providing precise quantitative insights. The outcomes reveal that the novel employs simple language and relatively straightforward sentences to enhance readability and comprehension, thus broadening its audience and impact while infusing authenticity into the narrative and aiding readers in understanding the characters' essential traits. This approach eschews complex, convoluted sentence structures that could obfuscate clarity and direct meaning. Additionally, the text exhibits lexical diversity, covering many themes with a higher type-token ratio (TTR), indicative of linguistic variety. Moreover, the frequent use of first-person singular pronouns is notable in The Handmaid's Tale. Regarding content word distribution, verbs predominate, followed by nouns, pronouns, adjectives, and adverbs. Keywords in the novel are categorized into proper nouns, thematic keywords, and stylistic indicators. Through analysis, the narrative initially portrays Gilead's hegemonic authority and women's struggles within its oppressive regime, followed by Offred's voyage of self-discovery and her resistance against male supremacy. "Community with a Shared Future for Mankind" should be built among countries, people, and even between humans and nature to develop and grow harmoniously and sustainably.

#### REFERENCES

- [1] Adil Jaafar, E., & Abdul-Salam Jasim, H. (2022). A corpus-based stylistic analysis of online suicide notes retrieved from Reddit. Cogent Arts & Humanities, 9(1), 2047434.
- [2] Atwood, Margaret. "When Afghanistan Was at Peace." In Writing with Intent: Essays, Reviews, Personal Prose--1983-2005, pp. 207, New York: Carroll & Graf, 2005.
- [3] Atwood, Margaret. "When Afghanistan Was at Peace." In Writing with Intent: Essays, Reviews, Personal Prose--1983-2005, pp. 207, New York: Carroll & Graf, 2005.
- [4] Atwood, Margaret. "Margaret Atwood on What 'The Handmaid's Tale' Means in the Age of Trump." The New York Times, 8 May 2020, p. 7.
- [5] Atwood, Margaret. Survival. [edition unavailable]. House of Anansi Press Inc, 2012. Web. 15 Oct. 2022.
- [6] Atwood, Margaret. The Handmaid's Tale. Knopf Doubleday Publishing Group, 1985. Retrieved from https://ia801807.us.archive.org/18/items/the-handmaids-tale\_pdf/the-handmaids-tale\_ForLibrary.pdf.
- [7] Atwood, Margaret. The Handmaid's Tale, translated by Chen Xiaowei. Shanghai: Shanghai Translation Publishing House, 2020, p. 83.
- [8] Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and short-term memory. Journal of Verbal Learning and Verbal Behavior, 14(6), 575–589.
- [9] Barad, Karen. Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning. Durham, NC: Duke University Press, 2007.
- [10] Bloom H (1994) The Western Canon: The Books and School of the Ages. New York: Harcourt.
- [11] Culpeper, J. (2009). Keyness: Words, parts-of-speech, and semantic categories in Shakespeare's Romeo and Juliet character-talk. International Journal of Corpus Linguistics, 14(1), 29–59.
- [12] Dam, Daný van, and Sara Polak. "Owning Gilead: Franchising Feminism through Margaret Atwood's The Handmaid's Tale and The Testaments." European Journal of English Studies, vol. 25, no. 2, May 2021, pp. 172–89.
- [13] Fischer-Starcke, B. (2010). Corpus linguistics in literary analysis: Jane Austen and her contemporaries. Continuum.
- [14] Fletcher, W. H. (2006). "Phrases in English" home. Retrieved from http://phrasesinenglish.org/index.html.
- [15] Freud, Sigmund. Mourning and Melancholia. The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XIV (1914-1916): On the History of the Psycho-Analytic Movement, Papers on Metapsychology and Other Works, 237-258, 1917.
- [16] Gámez Fuentes, María José, and Rebeca Maseda García. "Nostalgia and the Dialectics of Contemporary Feminisms in The Handmaid's Tale." Science Fiction Studies, vol. 48, no. 1, 2021, pp. 77–93.
- [17] Green, C. (2017). Introducing the Corpus of the Canon of Western Literature: A corpus for culturomics and stylistics. Language and Literature: International Journal of Stylistics, 26(4), 282–299.
- [18] Halliday, M.A.K (1994). An Introduction to Functional Grammar. (2nd edition). London: Edward Arnold.

- [19] Jadwin, Lisa. Margaret Atwood's The Handmaid's Tale (1985): Cultural and Historical Context. Blackwell Publishing, 2009.
- [20] Leech, G., & Short, M. (2007). Style in Fiction. Pearson Education Limited.
- [21] McIntyre, D., & Walker, B. (2019). Corpus stylistics: Theory and practices. Edinburgh University Press.
- [22] McNees, E. (2003). 'Tell': Narrative Technique in "The Handmaid's Tale." In S. Stein & C. Robin (Eds.), Approaches to Teaching Atwood's The Handmaid's Tale and Other Works (pp. 119-124). The Modern Language Association of America.
- [23] Revelles-Benavente, Beatriz. "Intra-Mat-Extuality: Feminist Resilience within Contemporary Literature." European Journal of English Studies, vol. 25, no. 2, May 2021, pp. 190–206.
- [24] Sanyal, J. (2006). English: The book for every englishspeaking Indian. Viva Books Ltd.
- [25] Savona, G. (2016). 'Historical Notes on The Handmaid's Tale: The 'The' and the Word 'of as Ambiguous Possessive Markers'. Partial Answers: Journal of Literature and the History of Ideas, 14(1), 15-39.
- [26] Schneider, Ana-Karina. "Clinging to Flesh: Embodied Experience in Contemporary Women's Dystopias." Contemporary Women's Writing, vol. 16, no. 3, Dec. 2022, pp. 343–60.
- [27] Scott, M. R. & Tribble (2006). Key Words and Corpus Analysis in Language Education. Amsterdam & Philadelphia: John Benjamins.
- [28] Scott, M. R. (1997). 'PC analysis of keywords and key keywords.' System, 25 (2), 233--45.
- [29] Scott, M. R. (2000). Focusing on the Text and Its Key Words, in L. Burnard & T. McEnery (Eds.), Rethinking Language Pedagogy from a Corpus Perspective, Volume 2. Frankfurt: Peter Lang., pp. 103--122.
- [30] Scott, M. (2014). WordSmith tools help. Lexical Analysis Software.
- [31] Scott, M. (2021). WordSmith tools [Computer software]. Lexical Analysis Software. https://www.lexically.net/wordsmith/downloads/.
- [32] Tolan, Fiona. Margaret Atwood: Feminism and Fiction. Rodopi, 2007.
- [33] Wang Kefei, & Hu Xianyao. (2008). A Corpus-Based Study of the Lexical Features of Chinese Translation. Chinese Translators Journal, 29(6), 6. Wang Kefei,&Hu Xianyao (2008) A Corpus based Study on the Characteristics of Chinese Vocabulary in Translation Chinese translation, 29(6), 6.]
- [34] Yuan Xia. Margaret Atwood: Queen of Canadian Literature. Wuhan: Huazhong University of Science and Technology Press, 2020. 34. [Yuan Xia. Margaret Atwood: Queen of Canadian Literature. Wuhan: Huazhong University of Science and Technology Press: 2020]
- [35] Qu, xing. (2021). Renlei mingyun gongtongti de jiazhi guan jichu. (4), 53-55. 35. [Qu Xing. (2021). The Value Foundations of a Community with a Shared Future for Humanity. (4), 53-55.]
- [36] Butler, Judith. The Force of Nonviolence: An Ethico-Political Bind. Verso, 2020.
- [37] Fletcher, William H. (2007). kfNgram (Version 1.3.14) [Computer Software]. Available at http://www.kwicfinder.com/kfNgram/kfNgramHelp.html.
- [38] Burgess, Anthony. (2024). novel. Britannica. Retrieved on Apr.13, 2024 from https://www.britannica.com/art/novel.
- [39] Anthony, L. (2022). TagAnt [Computer Software] (2.0.5). Waseda University. https://www.laurenceanthony.net/software/tagant/.
- [40] Scott, M.R. (2008). WordSmith Tools Help Manual. Version 5.0. Liverpool: Lexical Analysis Software.