| RESEARCH ARTICLE

How Ideology is reflected in *The Time Machine*: A Corpus-based Approach

Qian Liu

Postgraduate Student, School of Foreign Language, Jinan University, Guangzhou, Guangdong, China
Corresponding Author: Qian Liu, **E-mail**: liuqhic@163.com

| ABSTRACT

Much has been accumulated in the research on science fiction, corpus method to literary works, and critical discourse analysis on literary works, while research concerning the combination of these three elements is just beginning. The present study is a case study for examining how a corpus-based approach can combine with CDA and contribute to research on literary works. Specifically, Lancaster Semantic Analysis System (USAS) is firstly used to perform semantic encoding for the text of H.G. Wells' science fiction *The Time Machine*. Then the encoded text is imported into Sketch Engine, the ultimate tool to explore how language works. Second, the word list and the keyword program are used for word filtering. The filtered words are then divided into 3 categories, namely, character, environment, and psychology, according to different descriptive aspects. Third, the distribution and collocation of object words in different categories are tested by the sketch engine programs or USAS. Finally, CDA is carried out on these data in combination with the time of the text. Findings from the study have shown that language in *The Time Machine* is ideology-loaded, characterized by the distinctive modification of different characters, the vagueness of the psychological process, and the diversity of narrative perspectives. In response to scepticism of quantitative stylistics from literary critics, this paper serves to reinforce the literary value of simple quantitative text and corpus data. At the theoretical level, this study tries to explain how micro textual resources can interface with macro discourse, such as ideology and social cognition. At the methodological level, this study promotes the application of the combination of corpus linguistics and critical discourse analysis in stylistics.

| KEYWORDS

Corpus stylistics, critical discourse analysis, science fiction, *the time machine*

1. Introduction

In addition to being a novelist and historian, H. G. Wells was also a social reformer and a prophet. As a significant figure of the Fabian Society, a socialist group advocating democracy, he once wrote a number of masterpieces that focused on the present and reflected on the future. The first science fiction of H.G.Wells, *The Time Machine*, was published in 1895. It is futuristic science fiction of fewer than 50,000 words, and computer-readable versions are available (e.g., from Project Gutenberg). In this book, Wells railed against the iniquities of capitalism society with bizarre plots and exaggerations and predicted that science plays an important role in society but also brings terrible disasters to human beings. As a masterpiece in science fiction, the most representative feature of *The Time Machine* is its political nature, which plays an important role in predicting and enlightening human spiritual civilization, ecological civilization, political system, and social structure, reflecting the author's ideological connotation and the mainstream ideology of the society at that time. Existing studies on Wells and his works mainly focus on the translation and introduction in different ages and the enlightenment to social development (蔡立胜, 2006; 陈少娟, 2009; 豆瑞杰, 2010; 龙月圆, 2017; Keenan, 2018), most of which choose the qualitative methods advocated by literary criticism. However, Glaser and Strauss mentioned in 1967 that the sampling procedure in qualitative research is always flexible, ongoing, and evolving while selecting successive data (as cited in Khan, 2019). For this reason, this paper attempts to explore the linguistic features of Wells' *The Time Machine* using a corpus-based approach (梁茂成, 2010), which is combined with critical discourse analysis so as to uncover the ideology implied in this fiction.

1.1 Literature on Science Fiction

Science Fiction writings and movies have already been investigated by scholars around the world, including sociologists, anthropologists, linguists, and even historians (Hassler & Wilcox, 2008; Gendron, Ivanaj, Girard, & Arpin, 2017; Rolls & Rodgers, 2017; Araujo & Diniz, 2020). These works provide different accounts demonstrating the potential of science fiction in scientific research by imagining and illustrating social changes and recognizing the heuristic potential of science fiction explicitly. As Andrew Ross suggests, "Science fiction writers, more than those of any other pop genre, have been passionately concerned about their social responsibility to imagine better futures (as cited in Gerlach & Hamilton, 2003)". For this reason, science fiction is often a reflection of social issues, and the study of science fiction also concerns society. To comprehend its various incarnations, Gerlach and Hamilton (2003) have identified three types of social science fiction research, namely, research that draws upon science fiction as a way of teaching social theory, which attaches great importance to the analysis of social situations and contexts (Greenberg, Milstead, Olander, & Warrick, 1975; Gendron et al., 2017), researches that turned upon the production, circulation, and consumption of science fiction as a genre (刘畅, 张文红, 2019), studies that focused on science fiction as a force involved in the construction of modern and postmodern world-view (Keenan, 2018; 韩玉洁, 2019), thus becoming a way of thinking and knowing. The present study adopts the third perspective mentioned above. Certain science fiction texts (*The Time Machine*) are thoroughly examined to illustrate the social contexts and situations in which the fiction is created, as well as how ideology is embedded in the work.

In response to scepticism of quantitative stylistics from literary critics, this paper serves to reinforce the literary value of simple quantitative text and corpus data. In order to make stylistic analysis useful to literary criticism, it is necessary to aim at revealing previously neglected deeper meanings of the text, for stylistics is a kind of interdisciplinary which stresses both linguistic description and literary interpretation (Salama, 2011). Next, the two research methods used in this paper, namely, corpus-based method and critical discourse analysis, will be reviewed respectively in their application to literary works.

1.2 Literature on corpus-based method to literary genre

According to the extent to which the research relies on the corpus method, corpus linguistics can be divided into three subtypes (梁茂成, 2010), namely, corpus-driven methods, corpus-based methods, and corpus-supported methods. The first method advocates that the theory comes entirely from the corpus, and the researcher does not make any presupposition. The second method encourages researchers to put forward hypotheses based on their intuition or existing theories and then use the corpus to verify their hypotheses. The third method claims that researchers should put forward theories and obtain data support from corpora when explaining them. Since the first method is too idealistic and the third is over-selective, the second one appears to be the most common one in corpus linguistics. As a matter of fact, these three methods are not mutually exclusive in practical research (see Stubbs, 2005). Based on the background knowledge of *The Time Machine*, both the second and the third methods are adopted in the present study to investigate and verify the theme and ideology reflected in it.

Existing literature reveals that corpus research on literary style can be roughly divided into the following three perspectives (卢卫中, 夏云, 2010): The stylistic analysis based on the annotation of a particular linguistic feature in a corpus, the study of the theme of a work by means of linguistic expression and the research on authors' style.

The first way to study corpus stylistics (e.g., Stubbs, 2005) is to take language use as the starting point and focus on a particular aspect of language features, such as words, phrases, grammatical structures, and rhetorical devices. This approach aims at studying the stylistic functions and the contextual factors, which affect language use, such as the type of discourse, the style of writing, and the theme of the work. It mainly includes the exploration of the collocations and the representations of speech, writing, and thought. For example, Stubbs (2005) studied the linguistic style of Joseph Conrad's *Heart of Darkness* by word-frequency list, word distribution, text structure, and concordance, using the "prose" discourse in Brown, LOB, Frown, and FLOB as parallel reference corpus. This study focuses on vague expression, unreliable knowledge, text structure, and intertextuality, trying to find the relationship between narrative mode, theme, and language characteristics. Stubbs found that the theme of fiction is not only expressed by content and function words that express ambiguity and uncertainty but also by the repetition of certain types of phrases and collocations, such as noun phrases containing abstract nouns and adjectives with negative prefixes, which express the character Marlowe's psychological uncertainty about reality, appearance, location, etc. Such expressions highlight the conflict between the so-called civilized society Marlowe left and the so-called primitive society he encountered.

A second way of conducting corpus stylistic study is to focus on specific literary works. Unlike what has been mentioned above, the second approach takes the stylistic function or theme of the work as the entry point. Stylistic researchers in this field (Inaki & Okita, 2005; Pace-Sigge, 2018) pay attention to the embodiment of the special style of the work. They make objective and detailed descriptions of all relevant lexical and grammatical patterns of literary works by using corpus retrieval software and then verify their intuitions by statistical analysis and index concordance. For example, Inaki & Okita (2005) conducted a case study on the

transformation of Alice's role in two fairy tales, Alice in *Wonderland* and *Looking-Glass*, based on a small-corpus-based method. They analyzed the character image, Alice, by extracting the verbs of saying related to it and the modifier of the word "Alice" and supplemented it with a word list and concordance tool. The results show that the "Alice" in *Wonderland* plays a cautious and passive role as an uninvited guest, while the "Alice" in *Looking-Glass* presents herself as an optimistic and independent explorer. A corpus-based vocabulary survey (Pace-Sigge, 2018) focuses on the frequency, collocations, and semantic associations of words that may express the presence or absence of economic and power structures. Pace-Sigge investigated British literary works of the 19th and 20th centuries in this study, attempting to answer the question of whether the key concepts of Marx's critique of political economy have left a discernable influence on the literary language.

The third aspect of corpus stylistics concerns the specific style of a writer (e.g., Hoover, 2007; Wright, 2017). It mainly includes the study of copyright attribution, the diachronic study of style change, the forensic identification of style differences, representing the diversity of research topics, and the rigor and effectiveness of research methods. For example, Hoover (2007) adopted cluster analysis to explore the diachronic changes in Henry James' style and identified two different styles as well as five possible sub-styles. Wright (2017) sets out to address the forensic problem using a corpus linguistic approach and the 176-author 2.5 million-word Enron Email Corpus. Wright tests the accuracy of word n-grams in identifying the authors of anonymized email samples, drawing on literature positing the idiolectal nature of collocations, phrases, and word sequences.

In a nutshell, the aspects mentioned above take linguistic features, the theme of the studied works, and the style of the writer as the entry points, respectively. The current study is mainly unfolded from the second aspect, verifying the theme and ideology of the work through linguistic features such as word frequency and distribution, keywords, collocation, word bundle, and discourse prosody.

1.3 Literature on critical discourse analysis applied to literary works

Critical discourse analysis (CDA) is a linguistic method that emerged in England in the late 1970s. As a branch of linguistics, CDA involves the research of many disciplines such as psychology, sociology, stylistics, and mass communication (刘小策, 2017). It is often defined as instrumental in that language is regarded as more of social practice than a medium of communication in this field, aiming to explore the ideology hidden in language by analyzing the characteristics of language and its social and cultural background so as to reveal the complex relationship between language, power, and ideology (廖菲, 2017). In the definition of ideology, the question comes into the domain of philosophy. Ideology can be regarded as the understanding and cognitive process of common things. It is one kind of collection that contains all the ideas, opinions, concepts, thoughts, and values (龙月圆, 2017). Ideology is not the inherent thing in the human brain but from social existence. It can be affected by many factors, such as personal thinking ability, surrounding environment, education, and value orientation. Because of the various ideologies, there are many different understandings of the same thing. What CDA really cares about is the complex social environment in which language exists, rather than language units. Language is not only considered to be the reflection of social reality but also a part of social practice. CDA connects discourse and society and expands the critical perspective of discourse research, showing the great potential of discourse research from a critical perspective. The objects and methods of research in CDA are constantly developing and being innovated.

As for the objects of discourse research, CDA studies aim at political discourses, ranging from media discourse, racial discourse, and gender discourse to thematic discourse. Through the analysis of media language (e.g., Törnberg & Törnberg, 2016) or official documents (e.g., Kamasa, 2013), the public can be more sensitized to language, thus being able to recognize the delicate relationship between language and social life and to recognize the relationship between language and unequal rights. Typical sources of object texts are public speeches, book reviews, news reports, and government documents (Khan et al., 2019; 杨倩, 2019; Geertsema-Sligh & Worthington, 2020), while literature discourses such as novels (刘小策, 2017; 刘志宇, 2018) and script (廖菲, 2017) are also gaining attention. The researchers mentioned above have studied the ideological tendencies and power relationships of literary works, such as George Orwell's *Animal Farm*, Toni Morrison's *A Mercy*, and Quentin Tarantino's *The Hateful Eight*. However, compared to the typical sources, the scope of literary works in the CDA domain is relatively small. There are even fewer studies that combine the corpus method with CDA research (Salama, 2011; Wang, 2018).

As for the methods of research, just as Van Dijk (2001) argued, "CDA can be conducted in, and combined with any approach and sub-discipline in the humanities and the social sciences", Halliday's Systemic Functional Grammar (1973) and Fairclough's Three-Dimensional Modal (1995) are frequently applied as analytical tool and framework respectively in CDA (廖菲, 2017). Just as in other branches of discourse studies (Baker, 2006), CDA is always more specific and selective in order to choose the extracts from any speech, interview, or statement on the basis of keywords. In addition to the qualitative inquiry, researchers are not concerned with how much the sample is representative but is rather focused on understanding the subtle meanings within the phenomenon. In a more qualitative aspect, CDA research, combined with corpus technology, mainly identifies the intrinsic meaning of discourse by

observing word frequency, keyword concordance, collocation, word bundle, grammatical structure, and pattern. For example, Stubbs' research (1996) on Baden Powell's last report to the Boy Scouts and Girl Guides, which starts from analyzing their collocation distribution of the content words with the highest frequency, "happy" and "happiness", reveals the ideology hidden in single words and language patterns. An expression such as "Boys are told to be happy" and "Girls are told to make others happy" affect people's value orientation to a great extent. CDA is also carried out on the reports of migrant workers by commercial media and official media (Liu & Leung, 2021). Keywords analysis found that the negative words (e.g., discrimination, physical assault, etc.) about the survival of migrant workers in the two types of reports are significantly reduced, while the positive words (e.g., the increased income, the active entrepreneurial policy for migrant workers, etc.) are significantly increased. This phenomenon was then attributed to harmonious political ideology in China.

As we can see from 1.1, 1.2, and 1.3, much has been accumulated in the research on science fiction, corpus method to literary works, and critical discourse analysis on literary works, while research concerning the combination of these three elements is just beginning. The present study is a case study for examining how a corpus-based approach can combine with CDA and contribute to research on literary works, which is led by the following two questions:

1) How were the main characters and scenes in *The Time Machine* constructed?
2) What class stance and ideologies does H. G. Wells present in *The Time Machine*?

In response to skepticism of quantitative stylistics from literary critics, this paper serves to reinforce the literary value of simple quantitative text and corpus data. At the theoretical level, this study tries to explain how micro textual resources can interface with macro discourse, such as ideology and social cognition. At the methodological level, this study promotes the application of the combination of corpus linguistics and critical discourse analysis in stylistics.

2. Methodology

In order to uncover the ideological connotations implied in *The Time Machine* and arouse reflections on H. G. Wells' class stance, the present study uses corpus tools, namely, Sketch Engine (Kilgarriff, Rychly, Smrz, & Tugwell, 2004; Kilgarriff, Jakubicek, Kovar, Rychly, & Suchomel, 2014) and Lancaster Semantic Analysis System (USAS), to explore the linguistic characteristics of the novel and discover the narrative mode and theme of the novel. To be specific, Herbert George Wells' *The Time Machine* is firstly tagged by USAS, and the tagged text is then built into a small corpus using Sketch Engine. Taking English Web 2020 (enTenTen20) as our reference corpus, the word list and the keyword program are used for word filtering. In corpus linguistics, keyword analysis is employed for identifying prominent lexical items in the observing corpus as compared to the reference corpus. The most frequently occurred and meaningful words are chosen as the objects in the present study. As for the second step, since there are abundance of appearance, psychological and environmental descriptions in *The Time Machine*, which mean a lot to the story (龙月圆, 2017), the chosen words are then divided into 3 categories, namely, character, environment, and psychology, according to different descriptive aspects. Third, the distribution and collocation of object words in different categories are tested by the word sketch program and Lancaster Semantic Analysis System (USAS). Finally, in order to explore the author's ideology presented in his or her work, it is considered necessary to take into account the context of the author's life and personal experience; thus, an critical discourse analysis of these data is then conducted.

As a result, a total of 50 target words have been identified. Each of these words has then been classified into three groups (in bold) according to different descriptive aspects:

1) **Main characters** (*the time traveler, weena, eloi, morlocks,* etc.) and
2) **Psychological process** (*feel, think, seem, find, want,* etc.) and
3) **Environmental entities** (*Palace, under-world, upper-world,* etc.)

4. Findings and Discussion
4.1 Main characters in The Time Machine
As their very nature suggests, one of the reasons why counting individual words is not sufficient for corpus research is that the interesting content words are not evenly distributed across the text but are clustered at different places. For example, some words (e.g., *the time travellar*) occur only in the opening and closing narrative frames and are thus used to mark text structure (see Figure 1). The main part of the novel is a story told by the Time Traveller. However, when he tells his adventurous story, there are many audiences, including the medical man, the editor, the journalist, the psychologist, and "me". After finishing his story, there are also conversations between them. The Editor raises his doubtfulness about the story after the Time Traveller tells it, which is the question raised by all the readers. It gives Time Traveler a chance to convince the bystanders in fiction and reader of the fiction in the real world again and contributes to the diversity of narrative perspectives.
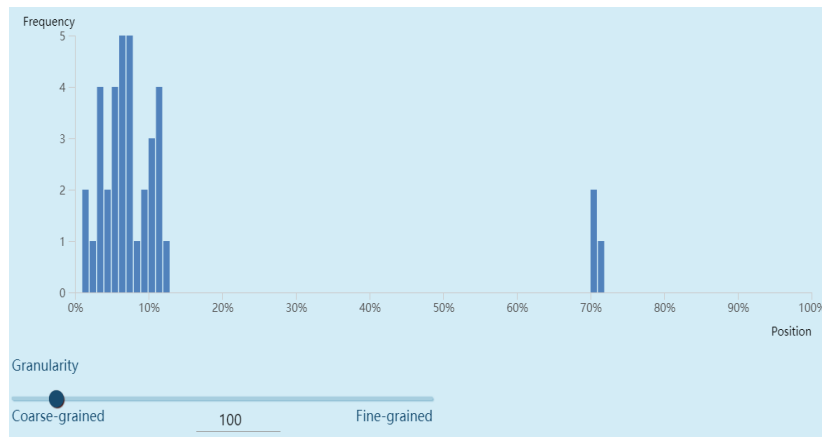
Figure 1 Distribution of *the time travellar* in the corpus

In *The Time Machine*, the future humans split into two distinct species: the limp, surface-dwelling Eloi and the ugly, nocturnal Morlocks. The Time Traveler believes that the Morlocks are descendants of the working class because of the large number of machines under the Eloi cities. On the other hand, Eloi are descendants of the leisure class of the time and seem to have lived a simple and luxurious life thanks to the efforts of the Morlocks. The word *Eloi* and *Morlocks* have a significant difference in frequency. The former occurred 8 times and the latter 35 times. This may be because the time traveler became friends with the Eloi Weena in the fiction, who then became the epitome of her racial image. Word sketches for *Eloi* and *Morlocks* (see Figure 2) show that Wells constructed the two characters quite differently.

In terms of the collocation of prepositions and the two races, the author uses the preposition "against" and "by" with *Morlocks* several times, building this race into enemies that need to be defeated by force, while the preposition "among" with *Eloi* is more static and peaceful.

- ...for nothing, I knew, would be more efficient **against** these Morlocks.
- I fancied that if I could solve their puzzles, I should find myself in possession of powers that might be of use **against** the Morlocks.
- I had in my possession a thing that was, perhaps, the best of all defenses **against** the Morlocks –I had matches!
- I now felt safe against being caught napping **by** the Morlocks...
- Then I remember Weena kissing my hands and ears and the voices of others **among** the Eloi.
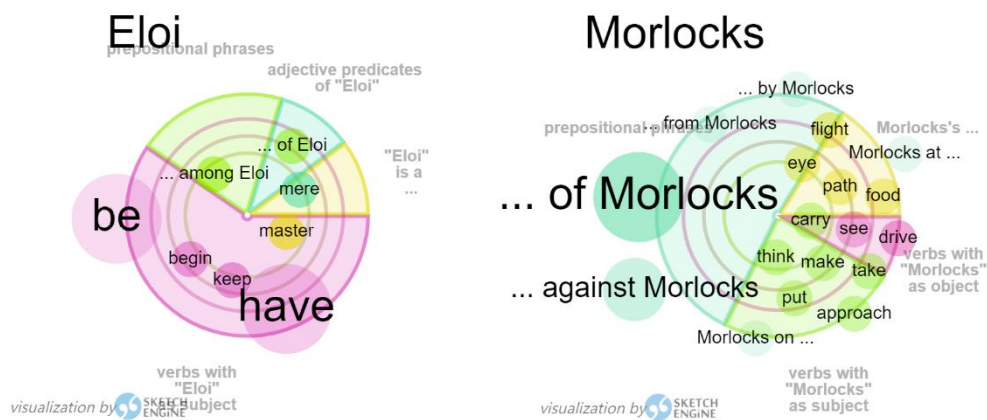

Figure 2 Word sketches for *Eloi* and *Morlocks*

In terms of the possessive case of the two races, *Eloi* goes with an abstract noun such as "influence", while *Morlocks* goes with a nonabstract noun such as "food", "eyes," and "path". This is mainly because the author constructs the Morlock people like a pragmatic race striving for survival, while the Eloi people as fatuous "monarch" living in a fantasy paradise.

- Probably my shrinking was largely due to the sympathetic **influence** of the Eloi, whose disgust of the Morlocks I now began to appreciate.
- The Morlocks' **food** had run short.
- The Morlocks' **eyes** shone like carbuncles.
- I followed in the Morlocks' **path**.

In terms of verbs with the two races as subjects, the *Morlocks* people enjoy a high degree of action complexity, while the verbs paired with *Eloi* are extremely monotonous (see Figure 3). Here, Wells constructs the *Morlocks* people as a race that takes the initiative to get what their want, while the *Eloi* as a race that seems to have a free ride but actually has no choices at all. The only thing Eloi can do is passively accept their fate.

- However great their intellectual degradation, the Eloi **had kept** too much of the human form not to claim my sympathy and make me perforce a sharer in their degradation and fear.
- These Eloi **were** mere fatted cattle, which the ant-like Morlocks preserved and preyed upon–probably saw to the breeding of.
- Why, too, if the Eloi **were** masters, could they not restore the machine to me?
- ...the Morlocks **had carried** my machine.
- So the Morlocks **thought**.
- Thrice I saw Morlocks **put** their heads down in a kind of agony and rush into the flames.
- I have suspected since that the Morlocks **had** even partially **taken** it to pieces while trying in their dim way to grasp its purpose.
- And the Morlocks **made** their garments,...
- But all was dark, and the Morlocks **had** their hands upon me.

Figure 3 verbs with "Eloi/Morlocks" as subject
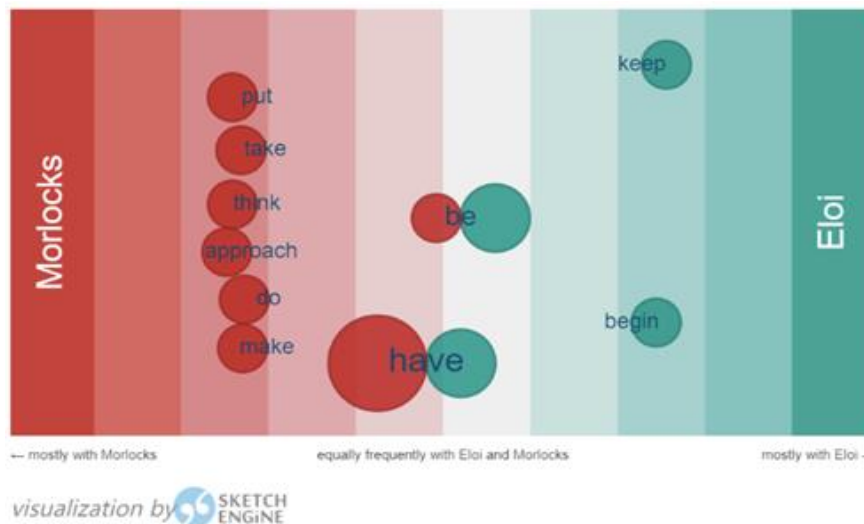
In *The Time Machine*, there is not only a huge and unbridgeable moral gap between different races but also the collapse of the moral value system within the same race. The Eloi, descendants of future humans, are a group of privileged and superior creatures. In a beautiful Eden free of strife and trouble, concepts of the value of life, virtue, and goodness, as well as affection and feelings, were superfluous in the Eloi world, a land of idleness, lack of ethics, and discipline. The loss of the purpose and needs of life makes Eloi people no longer think about the importance of virtue and ethics, which once existed above human beings and harmonized normal human life. It is the purposeless and meaningless nature of the Eloi life that causes the absence of human virtue in the

future utopian world. In Eloi's view, even the most important pursuit of virtue and good nature between people have been lost. There is no sincere affection left in them, only indifference and selfishness in human nature.
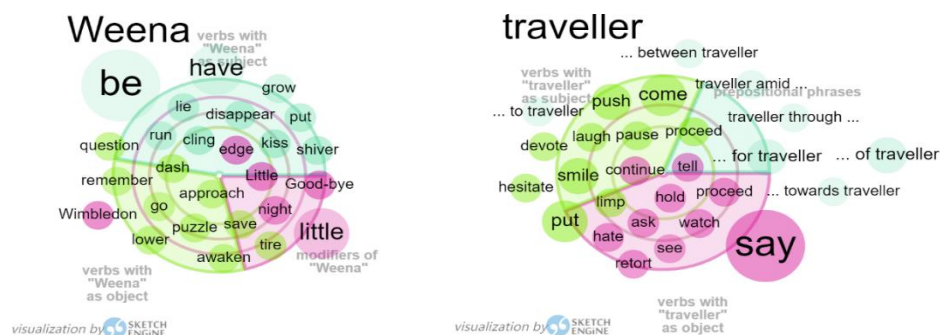

Figure 4 Word sketches for *Weena* and *traveller*

Another important character occurred in *The Time Machine* is an Eloi named *Weena*, which appears 54 times. No matter how time goes by and how the world changes, little Weena always follows the time traveler. The dangers and uncertainties of the future world did not change her trust and dependence on the time traveler. Unfortunately, all good things come to an end. After they enter the woods, they are surrounded by ugly and ferocious Morlocks, and little Weena is killed in a fire.

In terms of verbs of "Weena" as subject, the following verbs "cling", "shiver", "disappeared", "lay", "was gone," and "were lost" show some of the qualities Weena has as an Eloi, such as timid, cowardice, the lack of independent and always drift with the current and keep counting on others.

- It was very black, and Weena **clung** to me convulsively.
- And Weena **shivered** violently and became quite still.
- Everything save that little disk above was profoundly dark, and when I looked up again, Weena **had disappeared**.
- Then I turned to where Weena **lay** beside my iron mace.
- But Weena **was gone**.
- I knew that both I and Weena **were lost**, but I determined to make the Morlocks pay for their meat.

In terms of verbs of "Weena" as an object, the following verbs "lower", "approach," and "save" construct Weena either as a weak person who needs or is being rescued or as a person in desperate danger.

- I carefully **lowered** Weena from my shoulder as I halted and sat down upon the turf.
- I lit a match, and as I did so, two white forms that **had been approaching** Weena dashed hastily away.
- Some were bathing in exactly the place where I had **saved** Weena, and that suddenly gave me a keen stab of pain.

4.2 Psychological process in *The Time Machine*

Frequent nouns may indicate superficial topics in a text but not its underlying themes. Verbs are often a better candidate for stylistically relevant words (Stubbs, 2005). The Traveller is subject to the sort of "psychical uncertainty" (豆杰瑞, 2010). For this particular reason, Sketch Engine in the present study is used to lemmatize the text and list the top 10 psychological verb lemmas; they are:

- FEEL 108, THINK 100, SEEM 100, FIND 92, KNOW 80, WANT 30, UNDERSTAND 27, REMEMBER 24, BELIEVE 19, SUPPOSE 17

The Psychological lemma FEEL is the most frequently used one in this aspect of description (see figure 5). The results obtained from USAS show that both the words used to modify "feel" and the words modified by "feel" were overwhelmingly negative. The former (the first 3 examples) is used to emphasize how different the world of the future will be from the world of the present, while the latter (the other 8 examples) is used to construct the difficulties that the time traveler encountered in the future world as well as his uncertain mood.

- I **never** felt such a disappointment as I did in waiting five, ten, or fifteen minutes for an explosion that never came.
- I don't know if you will understand my feeling, but I **never** felt quite safe at my back.
- The serious people who took him seriously **never** felt quite sure of his deportment; they were somehow aware that trusting their reputations for judgment with him was like furnishing a nursery with egg-shell china.

Figure 5 Word sketche for "feel"

- I felt **hopelessly** cut off from my own kind–a strange animal in an unknown world.
- Presently I got up and came through the passage here, limping because my heel was still painful and feeling **sorely** begrimed.
- I felt **as perhaps** a bird may feel in the clear air, knowing the hawk wings above and will swoop.
- I felt **weary, stiff, and travel soiled**.
- I felt very **weary** after my exertion and sat down.
- It was hardly surprising then that as they drew nearer to Aldgate, the gateway to Whitechapel, he felt slightly **uneasy**.
- Very possibly, I had been feeling **desolate**.
- Somehow, this manner made me feel **ashamed** of myself.

The Psychological verb lemma KNOW (80 times) and UNDERSTAND (27 times) share two common points in *The Time Machine*. First, they are fairly evenly distributed throughout the text (see figure 6); Second, many instances are negative (see figure 7), either grammatically (I don't know, he did not know) or by implication (he wanted to know, you may hardly understand, you cannot know), which shows that this fiction is about the fallibility and distortions of human knowledge.
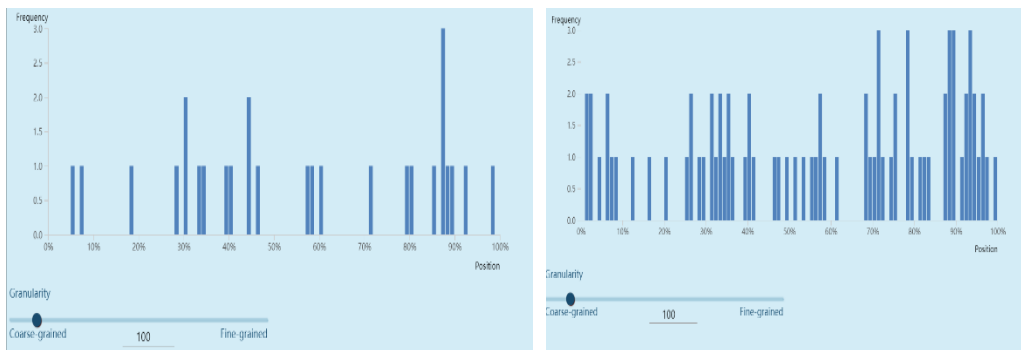

Figure 6 Distribution of *understand* and *know* in the corpus

- I do **not** know how long I sat peering down that well.
- The other men were Blank, the Editor aforementioned, a certain journalist, and another–a quiet, shy man with a beard– whom I **didn't** know.
- You **cannot** know how his expression followed the turns of his story!
- ...I tried to talk and found that her name was Weena, which though I **don't** know what it meant, somehow seemed appropriate enough.
- You may **hardly** understand how it took me.

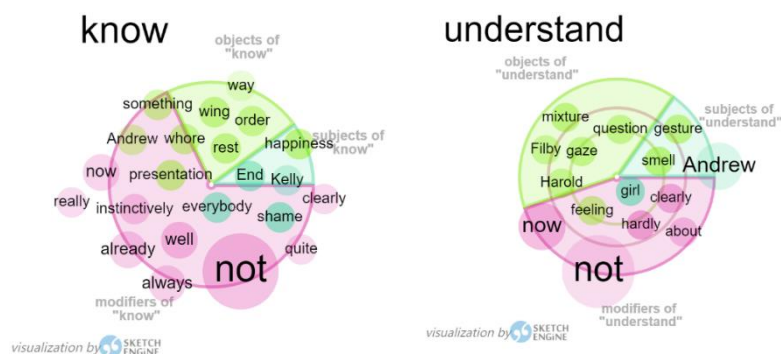- The wood, too, was full of a slumbrous murmur that I did **not** understand.

Figure 7 Word sketches for "know" and "understand"

4.3 Environmental description in The Time Machine

Since the time traveller comes to a new world and most of the readers will wonder what the future world will be like, Wells uses many words to depict the future world's surroundings in his work, and there are also many environmental descriptions in the original. Unlike Joseph Conrad's *Heart of Darkness*, in which major places are never named (Stubbs, 2005), Places usually mentioned in *The Time Machine* are *the under-world*, *upper-world*, and *the Palace of Green Porcelain*.
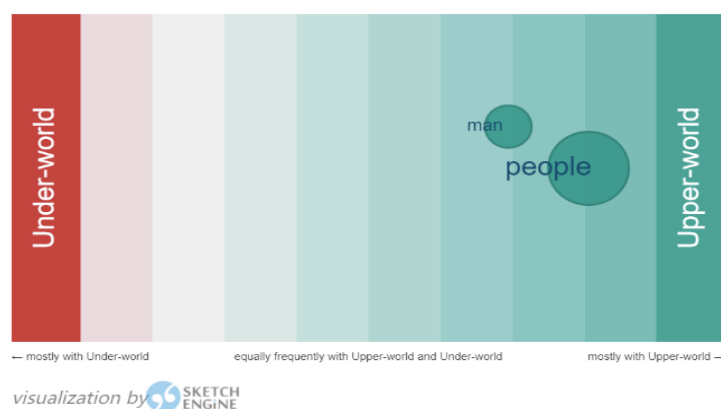

Figure 8 nouns modified by "Upper-world/Under-world"

From Figure 8 and subsequent concordance, we can find that Wells pays most attention to the people of the upper class since relevant nouns modified by *Upper-world* are "people", "man", "descendants," and "children", while *Under-world* usually goes with words such as artificial, feeding, industry. This collocation feature of hid language reflects Wells's petty-bourgeois thought.

- ...and as happy in their way as the Upper-world **people** were to theirs.
- And I now understood to some slight degree at least the reason for fear of the little Upper-world **people** for the dark.
- The Upper-world **people** might once have been the favoured aristocracy, and the Morlocks their mechanical servants:...
- ...that my graceful **children** of the Upper-world were not the sole **descendants** of our generation,...
- the Upper-world **man** had drifted towards his feeble prettiness...
- ...this **artificial** Under-world that such work as was necessary...
- ...went on, the **feeding** of the   Under-world, however...
- ...and the Under-world to mere mechanical **industry**...

Wells depicts the man of the Upper-world using the words"graceful children". What he wants to express is that the Elois are frail and naive. And for the other kind of humans, Wells says that they are "bleached, obscene nocturnal". All these words are derogative, which reflects that these men are not in a good image in the author's mind.

Other environmental images, such as the sun and flowers, are also an entry point to reflect the author's thoughts. Two strange white flowers, which represent the author's inner hope, and the sun, which indicates changes in the external environment, appear in *The Time Machine* over and over again. As the plot unfolds, both images become more and more negative, reflecting Wells' concern about the future of human society and his alarm about the political system at that time.

- The Time Traveller paused, put his hand into his pocket, and silently placed two **withered** flowers...
- His eye fell with a mute inquiry upon the **withered** white flowers upon the little table...
- As I did so, the shafts of the sun **smote** through the thunderstorm.
- The band of light that had indicated the sun had long since **disappeared**...
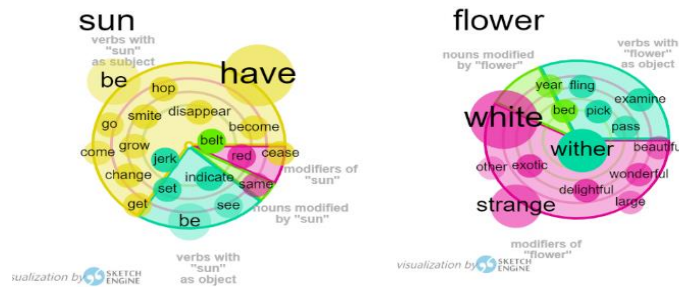

Figure 9 Word sketches for "sun" and "flower"

One thing is for sure: The age in which an author lived certainly influences the theme of his or her works. The first half of Wells' life was during the Victorian era, which is considered the culmination of Britain's industrial revolution and the heyday of the British economy and culture. One of the most striking features of the Victorian era was that Britain was one of the richest countries in the world. However, this wealth is more like an ultimate sin since the distribution of wealth has always been uneven in the Victorian era. The contrast between rich and poor is extremely obvious. Benjamin Disraeli, the famous Conservative Prime minister of this period, once said that Britain was a country of "two nations". On the one hand, there were nobles and landowners living luxuriously in their grand palaces; On the other, farmers and unemployed workers struggled to survive in their dilapidated huts. Just as Disraeli said: When the hut is uncomfortable, the palace is not safe. People's living standards are so different that there is a gap between heaven and hell in one country. The gap even became wider under the administration at that time. It should be admitted that the political system, relating to civil service, judiciary, parliamentary elections, and so on, established by the UK after the "Glorious Revolution" was the most advanced in the world at that time. However, during the period of the Industrial Revolution, with the rapid development of the economy and drastic changes in social structure, this system became more and more out of date and became a drop of poison that infected the whole tun of wine.

H. G. Wells was not only the author of novels such as *The Time Machine* and *The War of The Worlds* but also was a serious social and political thinker and educator. His influence was felt not only in science fiction but also in the improvement of human rights and global governance. In Victorian times, there was an idea that the survival of the fittest would explain the existence of different social classes. The Time Traveler speculates that Eloi are the result of the evolution of the nineteenth-century aristocracy, as well as the Morlocks descend from the working class. It is only when visiting the depths of Eloi city that the time traveler realizes the terrible truth of the simple balance between the Eloi and the Morlocks. The two different races have a certain antagonistic relationship with each other. Wells expresses his attitude towards the two races through the traveller's lens: his pity for the Eloi and his hatred of the Morlocks, which is a good confirmation of his petty-bourgeois position.

5. Conclusion

The present study attempts to examine how a corpus-based approach can combine with CDA and contribute to research on literary works through a case study on the science fiction *The Time Machine* completed by H. G. Wells. As corpus-based CDA integrates quantitative corpus linguistic methods, it allows researchers to scrutinize a broad range of discourse data and helps them maintain a high degree of objectivity when performing data analysis (Baker & McEnery, 2005). Findings from the study have shown that language in *The Time Machine* is ideology-loaded, characterized by the distinctive modification of different characters, the vagueness of the psychological process, and the diversity of narrative perspectives. The Eloi was constructed as a race that seems to have a free ride but actually has no choices at all. The only thing Eloi can do is passively accept their fate, either desperately waiting for rescue or kept in hopeless danger. On the other hand, the Morlocks people were described as a barbarous and vulgar race that took the initiative to get what they wanted. Descriptions of the environment are used to emphasize how different the world of the future will be from the world of the present and sentimentalize the uncertain and desperate emotion. At the same

time, Wells warns of the ecological problems that technology may bring. To solve the macroscopic environmental and ecological problems fundamentally, we must start from the spiritual and ecological problems of human beings. It can be said that social structure affects spiritual ecology and then environmental ecology. The vivid prophetic racial polarization in Wells' work reflects his concern about the ecological crisis and humanistic spirit.

It seems to be the duty of literary criticism to generalize the ideological and artistic features of the texts and to determine their diachronic or synchronic ideological and artistic value. Using concordance tools and large reference corpora to display and highlight relevant data, we can not only directly get important information related to the character and theme of literary works. It also provides reliable data support for us to interpret works, understand and digest or develop and question previous research results. Moving beyond pure qualitative analysis, which is the conventional way of critical discourse analysis, this study combines CDA with a corpus-based method, attempting to enrich the application of CDA in literary works and offer a means by which to account for literary critics. In response to scepticism of quantitative stylistics from literary critics, this paper also tries to reinforce the literary value of simple quantitative text and corpus data. At the theoretical level, this study tries to explain how micro textual resources can interface with macro discourse, such as ideology and social cognition. At the methodological level, this study promotes the application of the combination of corpus linguistics and critical discourse analysis in stylistics.

References
[1] Araujo, N., & Diniz, L. (2020). Relaciones entre la poesía de José Ángel Valente y la música espectral. *La Palabra,* (39), 71-84.
[2] Baker, P. (2006). *Using Corpora in Discourse Analysis*. London: Continuum.
[3] Baker, P., & McEnery, T. (2005). A corpus-based approach to discourses of refugees and asylum seekers in UN and newspaper texts. *J. Language Polit, 4*(2), 197–226.
[4] Burrows, J. F. (1987). *Computation into criticism*. Oxford: Clarendon Press.
[5] Disch, T. M. (1998). *The dreams our stuff is made of: How science fiction conquered the world*. New York: Free Press.
[6] Fairclough, N. (1995). *Critical Discourse Analysis: The Critical Study of Language*. London: Longman.
[7] Geertsema-Sligh, M., & Worthington, N. (2020). News Constructions of South Africa's Trial of the Century: Identity Discourse in the Steenkamp Shooting and Pistorius Trial, *Journalism Studies, 21*(15), 2196-2213.
[8] Gendron, C., Ivanaj, S., Girard, B., & Arpin, M. (2017). Science-fiction literature as inspiration for social theorizing within sustainability research. *Journal of Cleaner Production, 164*, 1553-1562.
[9] Gerlach, N., & Hamilton, S. N. (2003). Introduction: A history of social science fiction. *Science Fiction Studies, 30*(2), 161-173.
[10] Greenberg, M. H., Milstead, J. W., Olander, J. D., & Warrick, P. (Ed.). (1975). *Social problems through science fiction*. New York: St. Martin's.
[11] Halliday, M. A. K. (1973). *Exploration in the Function of Language*. London: Edward Arnold
[12] Hassler, D.M., & Wilcox, C. (2008). *New Boundaries in Political Science Fiction*. Columbia: Univ of South Carolina Press.
[13] Hoover, D. L. (2007). Corpus stylistics, stylometry, and the styles of Henry James. *Style, 41*(2), 174-203.
[14] Kamasa, V. (2013). Naming "In Vitro Fertilization": Critical Discourse Analysis of the Polish Catholic Church's Official Documents. *Procedia - Social and Behavioral Sciences, 95*, 154-159.
[15] Keenan, S. (2018). From historical chains to derivative futures: Title registries as time machines. *Social & Cultural Geography, 20*(3), 283-303.
[16] Khan,M. H., Adnan, H. M., Kaur, S., Khuhro, R. A.. Asghar, R., & Jabeen, S. (2019). Muslims' Representation in Donald Trump's Anti-Muslim-Islam Statement: A Critical Discourse Analysis. *Religions, 10*(2), 115-131.
[17] Kilgarriff, A., Rychly, P., Smrz, P., & Tugwell, D. (2004). The Sketch Engine. In *Proceedings from EURALEX 2004*, France: Lorient, 105-116.
[18] Kilgarriff, A., Jakubicek, M, Kovar, V., Rychly, P., & Suchomel, V. (2014). 'FindingTerms in Corpora for Many Languages with the Sketch Engine. In *Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics*. Gothenburg, Sweden: The Association for Computational Linguistics, 53–56.
[19] Liu, E., & Leung, J. H. C. (2021). Corpus insights into the harmonization of commercial media in China: News coverage of migrant worker issues as a case study. *Discourse, Context and Media, 41*, 100482.
[20] Pace-Sigge, M. (2018). How homo economicus is reflected in fiction - A corpus linguistic analysis of 19th and 20th-century capitalist societies. *Language Sciences, 70*, 103-117.
[21] Rolls, H., & Rodgers, M. P. H. (2017). Science-specific technical vocabulary in science fiction-fantasy texts: A case for 'language through literature'. *English for specific purposes, 48*, 44-56.
[22] Salama, A. H. Y. (2011). *Ideological collocation in meta-wahhabi discourse post-9/11: A symbiosis of critical discourse analysis and corpus analysis* (Doctoral dissertation). Lancaster University, UK. Retrieved from ProQuest Dissertations and Theses database.
[23] Stubbs, M. (1996). *Text and Corpus Analysis*. Oxford: Blackwell.
[24] Stubbs, M. (2005). Conrad in the computer: Examples of quantitative stylistic methods. *Language and Literature, 14*(1), 5-24.
[25] Tornberg, A., & Tornberg, Petter. (2016). Muslims in social media discourse: Combining topic modeling and critical discourse analysis. *Discourse, Context and Media, 13*, 132-142.
[26] Wright, D. (2017). Using word n-grams to identify authors and idiolects: A corpus approach to a forensic linguistic problem. *International journal of corpus linguistics, 22*(2), 212-241.
[27] 蔡立胜. (2006). 小说翻译创造性研究—以H. G. 威尔士的《时间机器》三个中译本为例[D]. 上海: 华东师范大学.
[28] 陈少娟. (2010). 赫·乔·威尔斯在中国的译介与影响[D]. 石家庄: 河北师范大学.

[29] 豆瑞杰. (2010). 从社会达尔文主义看赫伯特·乔治·威尔斯的《时间机器》. 石家庄: 河北师范大学.
[30] 韩玉洁. (2019). 论中国当代科幻小说生态向度的建设性. 小说评论, (3), 26-33.
[31] 廖菲. (2017).《八恶人》的批评话语分析[D]. 武汉: 华中师范大学.
[32] 刘畅, 张文红. (2019). 中国当代科幻小说出版研究. 出版广角, (24), 32-35.
[33] 刘志宇. (2018). 托尼·莫里森小说《恩惠》的批评话语分析[D]. 福州: 福州大学.
[34] 梁茂成. (2010). 理性主义、经验主义与语料库语言学[J]. 中国外语, 7(4), 90-97.
[35] 卢卫中, 夏云. (2010). 语料库文体学: 文学文体学研究的新途径[J]. 外国语, 33(1), 47-53.
[36] 刘小策. (2017).话语与权利—《动物农场》的批评话语分析[D]. 长春: 吉林大学.
[37] 龙月圆. (2017). 意识形态视角下清末民初科幻小说翻译策略研究——以The Time Machine杨心一汉译本为例[D]. 长沙: 湖南师范大学.
[38] 田海龙. (2006). 语篇研究的批评视角:从批评语言学到批评话语分析[J]. 山东外语教学, (2), 40-47.
[39] 杨帆. (2019). 论H.G.威尔斯科幻小说的生态反思[D]. 天津: 天津师范大学.
[40] 杨倩. (2019). 韩国网民对中国网络小说书评的批评话语分析[D]. 大连: 大连外国语大学.