
RESEARCH ARTICLE

The Study on Transitivity of Scientific Fiction: *The Time Machine* as an Example

Yuan Luo

Department of English, JiNan University, Guangzhou city, China

Corresponding Author: Yuan Luo, **E-mail:** 1114610116@qq.com

ABSTRACT

Corpus as a tool has been introduced in literature and linguistics for many years, allowing linguistic features and authorial style more visible and appreciable to readers. Based on Biber's MD/MF method and semantic field, this research is interested in linguistic features of science fiction and finds that transitivity is what makes science fiction distinguishable from science articles and novels. The current research is corpus-based, and corpora included in this research consist of *Silence Spring* (science article), *The Time Machine* (science fiction), and *Jane Eyre* (romance novel). This paper takes aims to investigate the 'transitivity' of science fiction by comparing three novels of different types. By means of the MD/MF method, the paper has found that transitivity in science fiction is achieved by TTR, mean word length, and mean sentence length. These findings may direct potential writers to focus on linguistic features worthy of notice, and readers can have a better understanding when they are reading novels.

KEYWORDS

Science fiction, linguistic features, MF/MD, Semantic field

ARTICLE DOI: 10.32996/ijllt.2022.5.6.14

1. Introduction

Science fiction, also known as Sci-Fi, is a form of literature combining features from science articles and novels, such as the authenticity of science articles and the fictionality of novels. H.G. Wells' *The Time Machine* makes a huge success and has been considered a pioneer in science fiction. This research takes an interest in the transitivity of science fiction and makes an attempt to know how science fiction constructs its transitivity in terms of linguistic features and semantic fields.

This research follows what methods Chen (2013) takes in *A corpus-based analysis of salient linguistic features of 'The Twilight Saga'*. In the linguistic aspect, type token ratio (TTR), mean word length, mean sentence length, and lexical sophistication will be examined in a bid to know how transitivity is exposed in terms of linguistic features listed above. The corpus tool used in this process includes AntConc for concordance and CLAWS for grammar tag. However, linguistic feature just reveals a partial truth about the transitivity of science fiction. To further prove the transitivity of science fiction, this research resorts to a semantic field represented by the E domain, which is achieved by USAS, a tool for semantic annotation. According to the results, science fiction shows a strong tendency of transitivity in both lexical and semantic aspects, and the transitivity in science fiction is realized by means of the lexicon, say lexical density and lexical sophistication, among others. Besides, the semantic field also plays an important role when it comes to transitivity; given the fact that E4 tops the main E-domains, joy and sorrow are reckoned as the major two emotions threaded through the whole text. Note should be made that joy and sorrow almost share an average level of distribution, showing that science novel is indicative of a kind of intermediacy in the aspect of emotion.

This current research is done with the great help of corpus, which has been widely used to conduct linguist or linguist-related research. In accordance with Chen(2013), Biber's MD/MF has a significant role to play in this paper. In addition, contrast has been employed to study the linguistic features of science fiction, say *The Time Machine*, in this research, and the texts chosen for comparison include the science article *Silent Spring* and the romance novel *Jane Eyre*.

1.1 Scientific Fiction

Science fiction, abbreviation SF or Sci-Fi, is a form of fiction that deals principally with the impact of actual or imagined science upon society or individuals. The term science fiction was popularized, if not invented, in the 1920s by one of the genre's principal advocates, the American Publisher Hugo, Gernsback. Science fiction is a modern genre. Though writers in antiquity sometimes dealt with themes common to modern science fiction, their stories made no attempt at scientific and technological plausibility, the feature that distinguishes science fiction from earlier speculative writings and other contemporary speculative genres such as fantasy and horror. The genre formally emerged in the West, where the social transformations wrought by the Industrial Revolution first led writers and intellectuals to extrapolate the future impact of technology. By the beginning of the 20th century, an array of standard science fiction "sets" had developed around certain themes, among them space travel, robots, alien beings, and time travel (see below Major science fiction themes). The customary "theatrics" of science fiction include prophetic warnings, utopian aspirations, elaborate scenarios for entirely imaginary worlds, titanic disasters, strange voyages, and political agitation of many extremist flavors, presented in the form of sermons, meditations, satires, allegories, and parodies—exhibiting every conceivable attitude toward the process of techno-social change, from cynical despair to cosmic bliss. Besides, science fiction films also draw attention from Babayi et al., aiming to examine how this language features in the movie in relation to other languages used and how it is rendered in the text by the makers of subtitles and by the actors. All these researches have revealed certain linguistic features of science fiction to some extent, while linguistic features of science fiction under the MD/MF approach are expected to be further investigated in the hope of having a deeper interpretation of how science fiction works.

The linguistic features of science fiction have aroused a lot of interest from various scholars. Podsievak et al. (2020:321-329) look at the linguistic features of science fiction from the perspective of cognitive linguistics through the approach of binary opposition. Theories in that study include linguistic, stylistic, and discursive analyses as well as cognitive linguistic analysis to ensure the reliability and validity of the obtained results. Olyanich and Rylshchikova (2016:237-244) investigate the role of lingua-semiotic creativity in science fiction discourse formation based on the creativity phenomenon in general and linguistic creativity in particular. Lingua-semiotic creativity is defined as a basic constitutive feature of science fiction discourse. Lingua-semiotic creativity is represented as the ability to handle images and put these images always as new signs, i.e., "to invent a new reality", then immersing them in the hyperreal (fantastic) discourse, the new communicative environment.

The Time Machine, the first novel by H. G. Wells, was published in book form in 1895. The novel is considered one of the earliest works of science fiction and the progenitor of the "time travel" subgenre. Wells advanced his social and political ideas in this narrative of a nameless Time Traveller who is hurtled into the year 802,701 by his elaborate ivory, crystal, and brass contraption. The world he finds is peopled by two races: the decadent Eloi, fluttery and useless, are dependent for food, clothing, and shelter on the simian subterranean Morlocks, who prey on them. The two races—whose names are borrowed from the biblical Eli and Moloch—symbolize Wells's vision of the eventual result of unchecked capitalism: a neurasthenic upper class that would eventually be devoured by a proletariat driven to the depths. The story follows a Victorian scientist who claims that he has invented a device that enables him to travel through time and has visited the future, arriving in the year 802,701 in what had once been London. There, he finds the future race, or, more accurately, races, because the human species has "evolved" into two distinct forms. Above ground live the Eloi—gentle, fairy-like, childish creatures whose existence appears to be free of struggle. However, another race of beings exists—the Morlocks, underground dwellers who, once subservient, now prey on the feeble, defenseless Eloi. By setting the action nearly a million years in the future, Wells was illustrating the Darwinian model of evolution by natural selection, "fast-forwarding" through the slow process of changes to species, the physical world, and the solar system. The novel is a class fable, as well as a scientific parable, in which the two societies of Wells's own period (the upper classes and the "lower orders") are recast as equally, though differently, "degenerate" beings. "Degeneration" is an evolution in reverse. While Wells's dystopic vision in *The Time Machine* is a deliberate debunking of the utopian fiction of the late nineteenth century, Wells represents a world in which the human struggle is doomed to failure.

This research aims at analyzing the linguistic features of science fiction represented by *The Time Machine*, through which we can get how science fiction is distinguished from romantic fiction and popular science.

1.2 MD/MF Method

The MF/MD approach is based on the assumption that statistical co-occurrence patterns reflect underlying shared communicative functions. It permits comparisons of multiple registers along a relatively small amount of underlying dimensions of variation. Dimensions (McEnery, Xiao & Tono 2006:161) represent 'different groupings of linguistic features that have been empirically determined to co-occur. With significant frequencies in texts'. To be specific, factor analysis, rather than a prior functional basis, is applied to identify the systematic co-occurrence patterns and group them into several distinct 'dimensions'. Then different texts and genres can be compared and interpreted in terms of communicative functions along each dimension.

The approach is therefore corpus-based, multi-dimensional, quantitative, and synthetic, namely combining quantitative techniques and functional interpretations. Originally, the MF/MD approach was designed to compare written and spoken registers in English. So far, the approach has been widely used in synchronic analyses of specific registers, diachronic research and register variation in other languages, and so forth.

According to Quaglio (2009:58), multi-feature/multidimensional (MF/MD) analysis is 'a quantitative corpus-based technique designed to find and interpret the co-occurrence of certain linguistic features in a corpus.' The availability of a corpus-based approach offers linguists the methodological tools required for multidimensional analysis. Some of the general characteristics of the MF/MD approach are as follows:

- (1) It is corpus-based, relying on analyses of a large number of natural texts, which should be attained with the aid of computational techniques. Given the drawbacks of an intuition-based approach to perceiving and analyzing variation in language uses, a corpus-based can provide a means of handling large amounts of languages and keeping track of many contextual factors at the same time.
- (2) It is multi-dimensional. Multiple parameters or dimensions of variation will be described in some registers.
- (3) It aims to analyze texts and registers instead of single linguistic constructions.
- (4) It synthesizes quantitative analyses and functional interpretation. Statistical techniques are used to identify and analyze co-occurrence patterns among linguistic features. Then these co-occurring linguistic features are interpreted in terms of communicative functions.

Considered as 'the most possible powerful tool for register and genre variation', Biber's MF/MD approach is established on the basis of Register. Fairclough (1995:14) defines register as "a socially ratified way of using language in connection with a particular type of social activity (e.g., interview, narrative, exposition)." In this definition, the language actually refers to linguistic features; the social ratified way refers to a functional interpretation of linguistic features; a particular type of social activity means a specific context for language use. Therefore, registers emphasize three essential elements: the linguistic features, the situational context, and the functional relationship between the former two components (Biber&Conrad,2009). Biber's MF/MD approach is based on the assumption that different registers would show different co-occurrence patterns of linguistic features. He referred to these sets of linguistic features that co-occur with high frequency as dimensions. Dimensions represent 'different classification of linguistic features that have been empirically confirmed to co-occur with markedly high frequencies' in registers. Composed of a set of co-occurring linguistic features, each dimension defines a different set of differences and similarities among various registers. Each dimension is postulated with varied interpretative labels, such as 'Narrative versus non-narrative concerns' and 'Explicit versus situation-dependent reference'.

In this research, MF/MD method will be used to analyze the linguistic features through five dimensions such as subjectivity vs. objectivity, narration vs. exposition, emotion vs. logic, complex grammar vs. simple grammar, and specific features vs. general features. For linguistic features, Type/Token Ratio, mean word length, mean sentence length, and lexical sophistication are put into examination in the following chapters.

2 Materials and Methods

2.1 Three Corpora

The corpora under study consist of three novels: *The Time Machine*, *Jane Eyre*, and *Silent Spring*. The time machine is science fiction with 32 901 words, *Jane Eyre*, a romance fiction with 189,951 words, and *Silent Spring* with 90,267 words.

The Time Machine is a science fiction novella by H. G. Wells, published in 1895. The work is generally credited with the popularization of the concept of time travel by using a vehicle or device to travel purposely and selectively forward or backward through time. The term "time machine", coined by Wells, is now almost universally used to refer to such a vehicle or device. Utilizing a frame story set in then-present Victorian England, Wells' text focuses on a recount of the otherwise anonymous Time Traveller's journey into the far future. A work of future history and speculative evolution, *Time Machine* is interpreted in modern times as a commentary on the increasing inequality and class divisions of Wells' era, which he projects as giving rise to two separate human species: the fair, childlike Eloi, and the savage, simian Morlocks, distant descendants of the contemporary upper and lower classes respectively. It is believed that Wells' depiction of the Eloi as a race living in plentitude and abandon was inspired by the utopic romance novel *News from Nowhere* (1890), though Wells' universe in the novel is notably more savage and brutal.

Jane Eyre is a novel by English writer Charlotte Brontë, published under the pen name "Currer Bell", on 16 October 1847, by Smith, Elder & Co. of London. The first American edition was published the following year by Harper & Brothers of New York. *Jane Eyre* is a Bildungsroman that follows the experiences of its eponymous heroine, including her growth to adulthood and her love for Mr. Rochester, the brooding master of Thornfield Hall. The novel revolutionized prose fiction by being the first to focus on its protagonist's moral and spiritual development through an intimate first-person narrative, where actions and events are coloured by a psychological intensity. Charlotte Brontë has been called the "first historian of the private consciousness" and the literary ancestor of writers like Proust and Joyce. The book contains elements of social criticism with a strong sense of Christian morality at its core, and it is considered by many to be ahead of its time because of Jane's individualistic character and how the novel approaches the topics of class, sexuality, religion, and feminism.

Silent Spring is an environmental science book by Rachel Carson. The book was published on September 27, 1962, documenting the adverse environmental effects caused by the indiscriminate use of pesticides. Carson accused the chemical industry of spreading disinformation and public officials of accepting the industry's marketing claims unquestioningly. Starting in the late 1950s, prior to the book's publication, Carson had focused her attention on environmental conservation, especially environmental problems that she believed were caused by synthetic pesticides. The result of her research was *Silent Spring*, which brought environmental concerns to the American public. The book was met with fierce opposition by chemical companies, but, owing to public opinion, it brought about numerous changes. It spurred a reversal in the United States' national pesticide policy, led to a nationwide ban on DDT for agricultural uses, and helped to inspire an environmental movement that led to the creation of the U.S. Environmental Protection Agency. In 1996, a follow-up book, *Beyond Silent Spring*, co-written by H.F. van Emden and David Peakall, was published. In 2006, *Silent Spring* was named one of the 25 greatest science books of all time by the editors of *Discover* magazine.

2.2 USAS for Semantic Field

The UCREL semantic analysis system is a framework for undertaking the automatic semantic analysis of the text. The framework has been designed and used across a number of research projects, and this page collects together various pointers to those projects and publications produced since 1990. The semantic tagset used by USAS was originally loosely based on Tom McArthur's Longman Lexicon of Contemporary English (McArthur, 1981). It has a multi-tier structure with 21 major discourse fields (shown here in table 2-1), subdivided, and with the possibility of further fine-grained subdivision in certain cases.

Table 2-1

A general and abstract terms	B the body and the individual	C arts and crafts	E emotion
F food and farming	G government and public	H architecture, housing and the home	I money and commerce in industry
K entertainment, sports and games	L life and living things	M movement, location, travel and transport	N numbers and measurement
O substances, materials, objects and equipment	P education	Q language and communication	S social actions, states and processes
T Time	W world and environment	X psychological actions, states and processes	Y science and technology
Z names and grammar			

In this research, semantic information about *The Time Machine* is examined by means of USAS for the semantic tag. For the sake of convenience, only the E domain would be chosen without checking all aspects of the twenty-one semantic domains listed above.

3. Results

3.1 Result of Lexicon

3.1.1 Lexical Density: TTR&STTR

Type/token ratio (TTR), also known as lexical density, is regarded as ‘the most common measure used in the investigation and of vocabulary scale and lexical richness’ (Fan 2011). Constrained by the size of corpora, TTR alone is not reliable for examining lexical diversity, and a Standardized type-token ratio (STTR) is thus introduced. According to Biber, TTR often co-occurs with other linguistic features such as mean word length and mean sentence length to show how diverse the vocabulary in a particular text is. Moreover, mean sentence length Basic statistics about lexical density in three corpora would be exhibited as follows:

Table 3-1 Basic comparison between TTM, JE, and SS

Code	The Time Machine	Jane Eyre	Silent Spring
tokens	32,901	189,951	90,267
types	4,675	12,949	8,862
TTR	14.21%	6.82%	9.91%
Standardized TTR	43.55%	45.73%	45.50%
Mean word length	4.3	4.16	4.87
Mean sentence length	17.18	18.61	21.14
1- letter word (%)	6.69	6.76	2.56
2- letter word (%)	16.13	18.09	17.6
3- letter word (%)	22.48	22.64	19.0
4- letter word (%)	17.92	18.23	15.4
5- letter word (%)	11.10	10.48	10.7
6- letter word (%)	7.81	7.87	8.6
7- letter word (%)	6.74	6.03	8.71
8- letter word (%)	4.43	4.04	5.79
9- letter word (%)	3.03	2.98	4.75
10- letter word (%)	2.00	1.46	3.33
11- letter word (%)	0.79	0.74	1.84
12- letter word (%)	0.60	0.40	1.07
13- letter word (%)	0.19	0.18	0.45
14+ -letter word (%)	0.10	0.09	0.24

(TTR=types/tokens*100%)

As the table shows, In *The Time Machine (TTM)*, TTR is 14.21% with 4675 types, much higher than *Jane Eyre (JE)* and *Silent Spring(SS)*, while the converse happens in terms of STTR. That happens because the size of *The Time Machine* is heavily smaller than the rest of the two books. Given that STTR is used to reduce the dependency of corpus size, here we take more consideration into STTR instead of TTR. As the result shows, in JE, STTR is 45.73% with 12,949 types, while that is 45.50% with 8,862 types and 43.55% with 4675 types, respectively, in SS and TTM, showing that the vocabulary in JE is more diverse than TTM and SS.

3.1.2 Word length & Sentence length

As can be seen from the above table, the mean word length of SS is 4.87, the highest one among the three corpora, whereas the mean word length in JE and TTM is separately 4.16 and 4.3, which indicates on average, SS tends to use longer words than JE and TTM. The table also presents the percentage of words comprising 1 letter to more than 14 letters. In SS, 82.57% of words are made up of words less than 7 letters (including 7 letters), and the percentage for TTM and JE is respectively 88.87% and 90.1%, which shows that SS uses fewer short words than the other two corpora. Conversely, SS has a higher percentage of long words comprising more than 14 letters when compared with the other two corpora. In sum, SS tends to use more complex words while JE seems to apply some shorter and simpler words, with TTM playing a middle role in this aspect.

In addition, the table also includes the average sentence length in the three corpora. The difference in mean sentence length is obvious: the mean sentence length of SS is 21.14, while JE and TTM are 18.61 and 17.18, respectively. Generally speaking, SS is apt to use longer sentences, and words comprising more than 14 letters are also more common to see than in JE and TTM, which can be explained by the clusters of proper nouns with a low frequency of occurrence. However, JE tends to use shorter words comprising of less than 7 words, and the average sentence length is much shorter than SS, ending up with better readability.

Notation should be paid to the fact that TTM possesses a sort of intermediacy between JE and SS, and that tribute also makes science fiction.

3.1.3 Content Words

Among various ways of classifying words, one way of them is to divide words into content words and function words. Content words include nouns, verbs, adjectives, and adverbs. These words "denote concepts such as objects, actions, attributes, and ideas that we can think about (Fromkin, Rodman & Hyams,2004). Function words include conjunctions, prepositions, and pronouns, which will be analyzed as part of some dimensions with the MF/MD method.

Table 3-2 Content words in TTM, JE, and SS(%)

Code	The Time Machine	Jane Eyre	Silent Spring
Noun (%)	18.60	8.57	25.67
Verb (%)	11.96	6.54	9.51
Adj.(%)	7.63	3.41	8.67
Adv. (%)	13.35	4.29	28.50
Total(%)	51.54	22.81	72.35

As the table suggests, the percent of content words in the three corpora is quite different. In TTM, nouns account for the highest percentage, 18.60%, and adverbs come the second with a percent of 13.35%. In JE, nouns take the highest percentage, 8.57%, and verbs lie in the second with a percentage of 6.54%. Lastly, in SS, adverbs take the highest percentage of 28.50%, followed by a noun which takes account for 25.67%. After looking at the data vertically, we should come to a horizontal view in a bid to get a whole picture. Among the three corpora, SS takes the highest account in nouns, adjectives, and adverbs with the percentage of 25.67%, 8.67%, and 28.5%, respectively, while TTM in verbs with the percentage of 11.96%. To sum up, three corpora in terms of the percentage of content words should be ranked as; SS>TTM>JE.

3.1.4 Lexical Sophistication

A keyword list is a tool in AntConc to figure out which words are unusually frequent(infrequent) in the corpus in comparison with words in a reference corpus. Besides, a Keyword list can allow you to know how many words (types and tokens) are comparatively salient in the context of wordlists consisting of words with different occurring frequencies, say most frequent 1000 words or 2nd 1000 most frequent words, which is achieved by means of keyness(log-likelihood) and effect. In this current research, three wordlists are extracted from BNC (British National Corpus) word list and named wordlist one, wordlist two, and wordlist three separately. Wordlist 1 consists of the most frequent 1000 words in BNC, wordlist 2 consists of the 2nd frequent 1000 words, and Wordlist 3 consists of uncommon words in BNC with the frequency of 100 times. By means of a Keyword list, we can make a comparison of the percentage of tokens and types between TTM, JE, and SS.

Table 3-3 Lexical frequency profile in TTM, JE, and SS

Code	Corpora	Tokens(%)	Types(%)
One	TTM	39.63	35.59
	JE	43.36	50.81
	SS	36.47	47.15
Two	TTM	91.75	41.04
	JE	89.37	35.93
	SS	86.52	33.49
Three	TTM	80.06	10.87
	JE	77.33	5.77
	SS	75.61	9.03

Table 3 allows us to see how the vocabulary of the three corpora distributes within the three wordlists or 3000 words. In wordlist one (the most frequent 1000 words in BNC), JE takes the top with tokens and types, respectively being 43.36% and 50.81%, followed by TTM and SS. In wordlist two (2nd frequent 1000 words in BNC), TTM takes the top with 91.75% tokens and 47.15% types, followed by JE and SS. As a result, in wordlist two, TTM also is put first in wordlist three, followed by JE and SS. That being said, the vocabulary of TTM concentrates on wordlist two and three, while that of JE and SS convene at word list one and word list two severally.

3.2 Result of Semantic Field

USAS, based on the Longman Lexicon of Contemporary English, is a tool used for semantic annotation according to the part of speech (pos), frequency, and context of words that need to be marked. There are two kinds of annotation, called the first annotation and second annotation; thus, more than one semantic annotation is available for each lexicon. In USAS, twenty-one semantic domains are included, with each domain incorporating its secondary domain. Like what scholars Xu et al.(Xu & Jiang& Zhan, 2021) have done with six translations of *LiSao*(one of the works by QuYuan), this research is also interested in the E domain((emotion) and its second domains as shown in Table 3-4.

Table 3-4 The semantic domains of E

E2	E2+	enthusiasm
	E2-	hatred
E3	E3+	calmness
	E3-	anger
E4	E4+	joy
	E4-	sorrow
E5	E5+	braveness
	E5-	terror
E6	E6+	confidence
	E6-	hesitation

The time machine consists of eighteen chapters, including the introduction and epilogue, the whole of which would be analyzed in terms of how the emotion is developed in the novel after the whole text is processed and retrieved by semantic annotation tool USAS and AntConc, a tool for concordance. The text being processed by USAS is like the following:

The_Z5 fire_O4.6+ burnt_O4.6+ brightly_O4.3,_PUNC and_Z5 the_Z5 soft_O4.5 radiance_O4.2+ of_Z5 the_Z5 incandescent_O4.2+ lights_W2 in_Z5 the_Z5 lilies_L3 of_Z5 silver_O1.1 caught_A9+ the_Z5 bubbles_O2 that_Z8 flashed_O4.3/N6 and_Z5 passed_M1 in_Z5 our_Z8 glasses_O2.

After this set of textual operations as described above, we now come to the following result:

Table 3-5 The distribution of E-domain

emotion chapters	E2		E3		E4		E5		E6	
	E2+	E2-	E3+	E3-	E4+	E4-	E5+	E5-	E6+	E6-
Intro	3	0	1	0	6	3	0	0	0	3
C-I	2	3	1	0	4	0	0	0	1	0
C-II	2	3	1	0	4	0	0	0	1	0
C-III	0	4	2	1	5	6	0	2	1	3
C-IV	0	1	3	9	6	3	1	5	1	0
C-V	4	0	1	0	11	2	0	3	1	0
C-VI	6	1	5	9	4	5	1	1	0	0
C-VII	1	0	3	7	9	8	1	12	1	5

(continued)

emotion chapters	E2		E3		E4		E5		E6	
	E2+	E2-	E3+	E3-	E4+	E4-	E5+	E5-	E6+	E6-
C-VIII	5	0	2	2	5	3	1	6	1	6
C-IX	5	3	3	10	3	2	4	8	2	1
C-X	7	2	1	2	5	1	15	0	2	5
C-XI	7	0	3	3	8	3	0	2	0	2
C-XII	0	0	11	0	3	4	1	8	0	1
C-XIII	1	0	3	4	5	2	0	1	1	1
C-XIV	2	0	3	6	0	2	0	3	0	3
C-XV	0	0	4	1	0	0	0	0	0	0
C-XVI	1	0	1	2	2	2	0	2	0	1
Epilogue	0	0	1	1	0	0	0	0	0	0
Total	46	17	49	57	80	73	24	53	12	31
	63		106		153		77		43	

(note: Intro stands for introduction, C for chapter)

As can be seen from the table above, the top of the table is five E-domains, while the vertical to the left shows seventeen chapters. It can be seen from the data in Table 3-5 how E-domain distributes throughout the whole novel: E2 occurs 63 times, E3 106 times, E4 153 times, E5 77 times, and E6 43 times. Closer inspection of the table displays that E2+ shows up 46 times while E2- for 17 times; E3+ shows up 49 times while E3- for 57 times; E4+ shows up 80 times while E4- for 73 times; E5+ shows up 24 times while E5- for 53 times; E6+ shows up for 12 times while E6- for 31 times. Comparing positive emotion domains with negative ones, we can find that positive emotion (211) is less common than negative emotion (231). Yet, it remains to be seen whether we can conclude that this novel is more negative than positive. After checking Table 3-4 in the above, E4 stands for joy(E4+) or sorrow (E4-), E5 for braveness(E5+) or terror(E5-), and E6 for confidence(E6+) or hesitation(E6-). What stands out in this table is that the frequency of E4 and E3 in the novel are significantly higher than in others, while E6 is the least frequent domain in the text. The difference between E2 +and E2-, E5+ and E5- is huge, while that between E3+ and E3-, E4+ and E4-, E6+ and E6- is relatively small. In E3, E3+ is close to E3- with a gap of only eight. Likewise, E4+ standing for joy and E4- standing for sorrow are almost close, with E4+ higher than E4- in E4. From this respect, *The Time Machine* is filled with elements of joy(E4+) and sorrow(E4-), anger(E3-), and terror(E5-), while hatred(E2-) and confidence(E6+) are less concerned through this novel.

Following a general examination of the result, we need to narrow it down into a more detailed report of these results from a look at how the E-domain develops among chapters. For the sake of convenience, this article only takes into consideration the most frequent items in E-domain when analyzing each chapter one by one.

Table 3-6 the E-domain of highest frequency in each chapter

	Fre- quency	In- tro	I	II	III	IV	V	VI	VII	VIII
E3+	3									
E3-	5					√		√		
E4+	7	√	√	√			√			
E4-	2				√					
E5+	1									
E5-	3								√	√
E6-	1									√

(continued)

	Frequency	IX	X	XI	XII	XIII	XIV	XV	XVI	EP
E3+	3				√			√		√
E3-	5	√					√		√	√
E4+	7			√		√			√	
E4-	2								√	
E5+	1		√							
E5-	3								√	
E6-	1									

Checking the most frequent E domains in each chapter, there are seven domains taking up the list: E3+, E3-, E4+, E4-, E5+, E5-, and E6-. It is clear that E4+(joy) is the most frequent domain distributed through seven chapters, followed by E3-(anger), E3+(calmness), and E5-(terror). Generally speaking, each chapter centers on only one domain, say E3+ or E3-, but the sixteenth chapter and epilogue are exceptions for Chapter XVI contain four domains while Epilogue occupies two domains.

4 Discussion

So far, in this research, we have looked at the novel *The Time Machine* from both lexical and semantic approaches. In the aspect of lexicon features of the novel, we have considered four items, say lexical density, word length & sentence length, content words, and lexical sophistication, and a comparison between three corpora in terms of the above four items. With regard to semantic dimension, we followed Xu et al. in coming to the semantic aspect by referring to the emotion domain using USAS for semantic annotation. Next, in the light of Biber's MF/MD methods, a closer look at the results and detailed analysis of them is available in the subsequent work.

4.1 Discussion of Lexicon

As shown in the previous part of 3.1, lexical density, word length & sentence length, content words, and lexical sophistication are combined together to serve the purpose of exposing the lexical features of science fiction. As shown in the result, TTM shows a sort of intermediacy between SS and JE with regard to lexical density, word & sentence length, and contents word, indicating that science fiction like TTM possesses both linguistic features from fiction like JE and popular science SS. According to Biber's MF/MD method, TTR, an index for lexical diversity, is a highly frequent index that often concurs with word length and acts as a negative factor. Next, we will analyze further the given result from the perspective of six dimensions in the subsequent texts.

4.1.1 Dimension 1: subjectivity vs. objectivity

In Dimension 1, the mean word length and TTR are said to show how objective and accurate a text is. To make it more concrete, that is to say, the bigger TTR and word mean length is, the less subjective or more objective a text is. Based on the data shown in the result, TTM is put in the middle between JE and SS in terms of mean word length and TTR; thus, it is reasonable to conclude that TTM is posited between objectiveness and subjectivity, creating a kind of half-authentic atmosphere for readers. That being said, science fiction takes the linguist features of both novel and popular science articles being subjective and objective separately. The novel is highly subjective due to its large amount of narration, description, and dialogue between characters. To highlight the personality of characters in the novel, some psych-verbs and emotional lexicons would also be included, adding more subjectivity to the novel. Note should be made that SS is closed to JE in this aspect because popular science article mainly consists of scientific stories and fiction.

4.1.2 Dimension 2: narration vs. exposition

Related closely to Dimension 1, Dimension 2 serves to tell the general narration from the professional exposition. In this dimension, mean word length reaches its maximum workload, and thus dimension 2 is evident enough to show that dimension 2 is where long words tend to show up. Furthermore, the grammatical function of dimension 2 lies in its obvious correlation to written form, highly dense information, independence to context, and clear expression of sense. As put by Biber in 1988 that longer words tend to express a more accurate and special meaning than what short words can convey; word length plays an especially fundamental role in dimension 2. To some extent, dimension 1 can be said to be a derivative of dimension 2 because it is an exposition that contributes to objectivity just as narration to subjectivity. Given that TTM shows a sort of transitivity between objectiveness and subjectivity as described in dimension 1, which is highly related to dimension 2 under discussion, TTM is believed to be both expository and narrative, which is a linguistic feature of science fiction integrating that of both fiction and popular science article. Besides, the novel is typical of narration, and this trait has been proved true before in various languages like English and Spanish by means of Multiple Dimension analysis. Contrastively, science article is characteristic of strong academy and technicality. As the grey belt between black and white, science fiction takes two sides at the same time.

4.1.3 Dimension 3: emotion vs. logic

Again, Mean word length is also employed in Dimension 3 as a negative linguistic feature for logic or emotion, indicating that mean word length is influenced by how much emotion or logic is conveyed in a text. Based on the result of content words shown before, three corpora should be ranked as SS>TTM>JE in terms of the percentage of content words. Content words are indicative of information and message, and a higher account of content words means less account of functional words like exclamations used to express emotion. In this respect, we can make a temporary conclusion that SS is the most informative among the three corpora; TTM is the second, followed by JE. That being said, the scale of three kinds of genres should be rated as popular science article, science fiction, and romantic fiction in terms of the amount of information they contain. Put in other words, fiction tends to express more emotion through more functional words, while science article is apt to convey information as much as possible through longer words, and logic is also indispensable in this process. For science fiction, logic and emotion are of the same importance because of their middle position between science articles and romantic fiction.

4.1.4 Dimension 4: complex grammar vs. simple grammar

In light of Biber's dimension, 4, mean word length together with lexical density (here we refer to TTR) can also function as an indicator of grammatical complexity. Mean word length and lexical density work together to reflect highly dense and contracted information. The novel shows a tendency of simplicity in grammatical aspects because it mainly focuses on narration and description. Whereas sentences in science articles contain more complex structures out of the need for logical argumentation. Thus, a possible conclusion could be made that science fiction shows a temperate nature of grammatical complexity, including a certain number of dependent sentences like cause clause, result clause, conditional clause, etc.

4.1.5 Dimension 5: specific vs. general features

Dimension 5 is about the specificity of the information conveyed in a text. Generally speaking, science article tends to contain more general information, while novel tends to be more specific in describing an event. In this dimension, quantifier, numeral, monosyllable, demonstrative pronoun, and others show up as positive features, while adjectives as adverbial modifiers, TTR and mean word length, etc., play a role in negative features in this aspect. Given the result based on TTR and mean word length, science fiction would express specific and detailed information by means of quantifier, numeral, monosyllable, demonstrative pronoun, first personal pronoun, and private verb; when it comes to general information, it would use an adjective as adverbial clause, disyllable, higher TTR and longer mean word length in a bid to show qualitative and categorical features. Note should be made that the function of adjective changes as its focus shifts from a tribute to action.

To sum up, in this chapter, we analyze the result in light of Biber's MD/MF method, where five features have been taken into consideration to look at linguistic features of science fiction as one the type of novel. This analysis is based on the data from results in the last chapter, and a table is necessary here to show the comparison between three corpora in terms of 5 dimensions.

	content	rank
Dimension 1	subjectivity vs. objectivity	JE>TTM>SS
Dimension 2	narration vs. exposition	JE>TTM>SS
Dimension 3	emotion vs. logic	JE>TTM>SS
Dimension 4	complex grammar vs. simple grammar	JE>TTM>SS
Dimension 5	specific vs. general features	JE>TTM>SS

Table 4-1 comparison between JE, TTM, and SS.

From the table above, we can get the information below:

- (1) In Dimension 1, science fiction shows a transitory trait between subjectivity and objectivity. For subjectivity, adverbs, the first person pronoun, exaggeration, intensifier, etc., would be put into use as positive features to show subjectivity in science fiction. However, nouns, TTR, and attributive words can be used as negative features to indicate objectivity.
- (2) In Dimension 2, science fiction takes the feature of both narration and exposition. Narration is demonstrated by auxiliary words, state words, and adjectives, while exposition is mainly achieved by mean word length and modal verbs.
- (3) In Dimension 3, science fiction stands on both sides of emotion and logic. Positive features would be such features as a modal particle, special question, second-person pronoun, exclamatory sentence, whereas negative features would be represented as link words.
- (4) In Dimension 4, science fiction is in the middle in terms of grammar complexity, which can be measured by the ratio of subordinate sentences in the whole text. In contrast, TTR and mean word length serve as an indicator for less complexity in the aspect of grammar or relative simplicity in grammar. That being said, lexical density and longer words can be used as alternants for complex sentences.
- (5) In Dimension 5, science fiction conveys both general and specific information. Specific information often takes the feature of a quantifier, numerals, while general information often applies adjectives for adverbial use.

So far, we've discussed the linguistic features of science fiction from the perspective of Biber's MD/MF method, and we can conclude that science fiction possesses a quality of transition between science articles and romantic fiction.

4.2 Discussion about Semantic Field

The result from the semantic field tells us that E3-, E4+, and E4- top the seven E domains, showing that anger, joy, and sorrow are three key emotions threading through the whole novel. For further analysis, the Wordcloud of each chapter has been made for a more objective and clear interpretation of how the three key emotion domains pioneer through the whole text, say *The Time*

Machine. A Word Cloud is a visual representation of some text as a bunch of words based on a weight associated with each word. Typically, the frequency (keyword density) of the words on the page is used as the weight even though some other metric might be used. *Word it out* is an online free word cloud generator that uses word frequency as the weight. Hence, the generated word cloud of a page gives a quick understanding of how the page is optimized for certain words. *Word it out* is also smart enough to extract phrases and works much better than most other word cloud generators on the web. The word cloud of a blog, news page, or feed will give a quick idea about what topics are being discussed. Since the novel contains as many as seventeen chapters, one way to simplify the current analysis is by breaking the whole novel into several parts with a common emotion domain or feature, whose wordcloud picture would be presented thereby.

To draw a wordcloud, a key wordlist is necessary. The key wordlists of each a chapter from introduction to epilogue would be given as below:

Table 4-2

Intro	Time said dimensions Man Traveller all Filby Space about move one any some three Psychologist line think only travel know the Fourth moment another along with real people back Medical Dimension other.
C-2	Time said Traveller machine Psychologist Man table through Medical upon all Look here traveled lever lamp hand future about thing took light sat one little mode lot her two some.
C-3	Time said Traveller Man Editor all dinner face story Journalist door Medical came towards Silent say again looked tell went back some most about traveling round word glass saw suppose
C-4	Saw like seemed machine one through minute again now upon came white time round took Presently over stood moment little more felt faint hail sun myself laboratory lever grew day.
C-5	Little people all flowers Time upon one saw came some white Machine more strange made perhaps felt thing fruit presently most fruits began moment seemed come flinging like mind tables.
C-6	All time now some little even conditions things upon first world here found saw under people well still physical great thought security more energy sunset strange about One Nature abundant.
C-7	Machine little came time all some sphinx myself people bronze lawn about bushes felt pedestal thought went round the world made face upon like way again ran Suppose thing white towards
C-8	One little time first white see some came saw Yet all made upon about great people even once world dark eyes thing seemed well like same enough again things though
C-9	Little well one light saw last darkness over match only even feeling time face Morlocks Weena again way felt without came another shaft all now dark handback hands great started.
C-10	Morlocks night Weena time some all now upon came old man Fear thought new felt human even like little before myself mind seen went seemed clear under great against black
C-11	Gallery upon found all time dust came Weena some side one little part thing went only thick place Here more first still once any about saw Palace even thought never
C-12	fire Weena Morlocks came seemed upon little wood some camphor felt still through one about behind before darkness now full back began again went red white saw forest black lit bar
C-13	Same one came upon thing towards hand perfect Time levers Underworld thought now Morlocks myself last little Sphinx habit pleasant all world bronze needs only Machine intelligence like change against
C-14	Sky saw the more red sun one sea upon grew slowly over time white green all now still seemed machine like again same beach earth black another moving hand air moon.
C-15	Saw again laboratory came thing door machine time passed old back seemed upon now across before slower like little Presently became familiar opened rest exactly Just exact golden brain ones.

C-16	Time said Traveller one upon hand looked Medical Man Machine door story came lamp took thing all put room here laboratory dream now come Editor believe pipe corridor see went.
Epilogue	Own still time among even men One part back Age mutual vast casual man's reef witness upon saline theory saw answered plesiosaurus-haunted say now days comfort lived memory ages wonder

The top thirty keywords have been taken into consideration when looking at the semantic aspect of the novel. The wordlist is made in accordance with the frequency of words occurring in the text. It's apparent to see that the time set in the novel is past tense, the topic is concerned with time travel, and the first person pronoun is frequently employed throughout the novel.

5 Conclusion

This research has investigated the science fiction *The Time Machine* from aspects of linguistic features and semantic fields. For linguistic features, Biber's MF/MD method has been employed to examine the whole text in terms of TTR, mean word/sentence length, and lexical sophistication. By comparing the linguistic features of *The Time Machine* (TTM), *Jane Eyre* (JE), and *Silent Spring* (SS), this research finds that *The Time Machine* or science fiction possesses a kind of intermediacy between romance fiction and science article. In this research, five dimensions have been included that is subjectivity vs. objectivity, narration vs. exposition, emotion vs. logic, complex grammar vs. simple grammar, and specific features vs. general features. In Dimension 1, science fiction shows a transitory trait between subjectivity and objectivity. For subjectivity, adverbs, the first person pronoun, exaggeration, intensifier, etc., would be put into use as positive features to show subjectivity in science fiction. However, nouns, TTR, and attributive words can be used as negative features to indicate objectivity. In Dimension 2, science fiction takes the feature of both narration and exposition. Narration is demonstrated by auxiliary words, state words, and adjectives, while exposition is mainly achieved by mean word length and modal verbs. In Dimension 3, science fiction stands on both sides of emotion and logic. Positive features would be such features as a modal particle, special question, second-person pronoun, and exclamatory sentence, whereas negative features would be represented as link words. In Dimension 4, science fiction is in the middle in terms of grammar complexity, which can be measured by the ratio of subordinate sentences in the whole text. In contrast, TTR and mean word length serve as an indicator for less complexity in the aspect of grammar or relative simplicity in grammar. That being said, lexical density and longer words can be used as alternants for complex sentences. In Dimension 5, science fiction conveys both general and specific information. Specific information often takes the feature of quantifier, and numerals, while general information often applies adjectives for adverbial use.

For the semantic field, the E domain is selected as the target domain for semantic analysis of *The Time Machine*, including five major domains from E2 to E5, with two secondary domains attached with a minus or plus subordinating to five major domains: E2+ enthusiasm vs. E2-hatred, E3+calmness vs. E3-anger, E4+joy vs. E4- sorrow, E5+ braveness vs. E5-terror, E6+confidence vs. E6-hesitation. The result from the semantic field tells us that E3-, E4+, and E4- top the seven E domains, showing that anger, joy, and sorrow are three key emotions threading through the whole novel. Besides, the keyword list also plays a role in deciding the semantic aspect of the novel. The top thirty keywords have been taken into consideration when looking at the semantic aspect of the novel. The wordlist is made in accordance with the frequency of words occurring in the text. It's apparent to see that the time set in the novel is past tense, the topic is concerned with time travel, and the first-person pronoun is frequently employed throughout the novel.

Based on corpus and MF/MD method, this research has investigated the transitivity of science fiction. The transitivity is achieved by showing linguistic features such as the middle level of TTR, mean word length, and sentence length. On the other hand, the semantic field also contributes to the transitivity of science fiction. To some extent, the current research serves as a support for the transitivity of science fiction. Still, there are some limitations shown in the following aspects: first of all, the linguistic features under discussion fail to cover all linguistic features discovered in Biber's framework. Secondly, the texts chosen as reference corpora are relatively random without full consideration of their size of them. Therefore, further research is suggested to have a closer look at other linguistic features other than TTR, mean word length, mean sentence length, and lexical sophistication, and the size of corpora should be expanded to cover various types of text as widely as possible.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References.

- [1] Biber, D. (1988) *Variation Across Speech and Writing* [M].Cambridge : *Cambridge University Press*.
- [2] Biber, D.(2003) Lexical bundles in academic speech and writing[J]. *Practical Applications in Language and Computers*.
- [3] Chen, Y. (2013) A corpus-based analysis of salient linguistic features of the twilight saga [D]. *Guang Dong University of Foreign Studies*.
- [4] Fan, X.M. (2011) Xiaoxue Yingyu Shixi Jiaoshi Ketang Huayu Leicibi Fenxi (An Analysis of Type-token Ratio of Teacher Talk in EFL Student - teaching Classroom)[J]. *Journal of Language and Literature Studies*(166).
- [5] Fromkin, V., Rodman, R.,& Hyams, N. (2004) *An introduction to language: Seven Edition*[M].*Beijing: Peking University Press*.
- [6] Liu Y. (2019) The orientation and characteristics of novel and other four styles in style variation patterns[J]*Jiangnan Academic*.38(1).
- [7] Parkinson, B. (1995) *Ideas and Realities of Emotion* [M].*London & New York: Routledge*.
- [8] Xu Z, Jiang Y &Zhan J. (2021) Translation and Translation of LiSao: Emotional Elements and Reconstruction of Qu Yuan's Image.[J].*Foreign Languages Research*.4(188).