Study on Stylistic Effects of Three-Word Clusters in *The Great Gatsby* from the Perspective of Corpus Stylistics

Qiyu Sun*

School of Foreign Languages, Shanghai Jiao Tong University, No.800 Dongchuan Road, Minhang District, Shanghai, China
Corresponding Author: Qiyu Sun, E-mail: sunqiyu@sjtu.edu.cn

ARTICLE INFO	ABSTRACT
	This paper examines high-frequency three-word clusters and their stylistic effects in Fitzgerald's *The Great Gatsby* through the analysis of the five types of clusters by means of corpus, providing a new perspective for the appreciation of the masterpiece *The Great Gatsby*. According to the analysis, we found that (i) Labels take up too small a percentage to reveal certain stylistic features in this novel. (ii) Speech clusters, occupying the largest percentage, are used to show characters' mental process and indicate interactions between the narrator and readers. Large amount of negation elements attached to clusters in this group help to deepen the tragic themes, i.e. the failure of American dreams. (iii) *As If X* clusters serve to portray characters in an objective way and disclose Gatsby's distorted inner world. (iv) Body Part clusters demonstrate Nick's inclination to observe others, which enhances the credibility of his narration. (v) Time clusters denoting short time span contribute to the compact storyline. The results demonstrated that high-frequency three-word clusters are of great significance to the interpretation of *The Great Gatsby*, which provides a new perspective to the analysis of literary works. **Key words:** *The Great Gatsby*, Corpus Stylistics, Three-word Clusters, Stylistic Effects

INTRODUCTION

At present, scholars have not reached an agreement on the definition of *cluster*, which results in up to 40 relevant terms (Wary 2002: 9), encompassing formulaic language, lexical phrases, lexical bundles, phraseological units, prefabs, etc.. Although clusters are labeled with different names in different contexts, *cluster* is preferably adopted in this thesis. Scott (2004) views cluster as "a group of words which follow each other in a text", which is adopted as the working definition in the current study for it's the most popular one among all the studies on clusters in literary works .

Studies on clusters have for decades received much attention from researchers . At present, however, studies on clusters are mainly concerned with language processing (Anna et al. 2011, Tremblay et al. 2011), language teaching (Meunier & Gouverneur 2007, Allan 2016), second language acquisition (Nekrasova 2009, Qin 2014, Huang 2015) and English for Special Purposes (ESP) (Grabowski 2015, Wang 2017). Nevertheless, there have been much fewer investigations of clusters in literary works (like Hardy 2003; Stubbs 2005), especially in recent ten years (Fischerstarcke 2009).

Additionally, almost all studies on clusters in literary works classify clusters according to their lengths, i.e. three-, four-, five- and six-word clusters, or according to their grammatical structures, instead of their functions. To better account for the textual functions of clusters, however, the current study attempts to adopt Mahlberg's (2007) classification of clusters based on their functions: Labels, Speech clusters, *As If X* clusters, Body Part clusters and Time and Place clusters.

Furthermore, as Mahlberg (2007) regards corpus stylistics as "a way of bringing the study of language and literature close together", to study literary works from perspective of corpus stylistics has now become a significant research paradigm with the rapid development of corpus linguistics. Aiming to analyze the stylistic features in literary texts objectively and quantitatively by means of corpus techniques and statistical methods, corpus stylistics provides an entirely new perspective for linguistic analysis of literary works. With the aid of corpora, some valuable stylistic information hidden in literary texts can be disclosed and therefore it makes useful contributions to literary criticism and literary research, as is put by Cao (2017). As a result, corpus is widely used in studies of literary works (Tabata 2002, Mahlberg 2010). Nonetheless, it has not been employed to explore *The Great Gatsby*, not to mention the study on three-word clusters in this novel from the perspective of corpus stylistics.

In response to existing limitations in research on both *The Great Gatsby* and clusters, the present study seeks to incorporate clusters into the analysis of literary works by investigating the role of high-frequency three-word clusters in Fitzgerald's novel *The Great Gatsby* from the perspective of

corpus stylistics in the hope of taking a comprehensive look at how theories of clusters can be applied to the analysis of literary works, so that stylistic effects will be acquired as to what roles high-frequency three-word clusters that are classified into five categories play in *The Great Gatsby*, hence serving to improve people's understanding and perception of the fiction. Additionally, this paper hopes to provide a new perspective for the appreciation of the masterpiece *The Great Gatsby*, via exploring the role of three-word clusters in the text. More briefly, the present study attempts to address the following five questions in *The Great Gatsby*: 1) What are the frequencies and stylistic effects of three-word Labels? 2) What are the frequencies and stylistic effects of three-word Speech clusters ? 3) What are the frequencies and stylistic effects of three-word *As If X* clusters ? 4) What are the frequencies and stylistic effects of three-word Body Part clusters? 5) What are the frequencies and stylistic effects of three-word Time and Place clusters?

RESEARCH METHODOLOGY

Quantitative and qualitative analyses were combined to carry out the investigation of those clusters in the present study. Particularly, quantitative methods were primarily adopted in the corpus-based measurement of clusters in *The Great Gatsby*, whereas qualitative approaches involved the observation, illustration and interpretation of the linguistic phenomena revealed by those clusters.

Corpus Source: The Great Gatsby

Francis Scott Key Fitzgerald is commonly regarded as a literary giant of America in the late 20th century and one of the most outstanding writers of his period. Renowned for his realistic views, satiric depiction of stories together with vivid portrait of "the hollowness of the American worship of riches and the unending American dreams of love, splendor, and gratified desires" (Zuo 2006: 320), Fitzgerald has produced numerous writings of lasting value, among which *The Great Gatsby* "is considered by many to be Fitzgerald's best novel" (Zuo 2006: 321).

The Great Gatsby, as a severe criticism of the whole society, is set in New York and Long Island. The young Mid-Westerner Jay Gatsby loves the beautiful Daisy, who has married the brute man Tom Buchanan for his wealth. In later years Gatsby, achieving fabulous material success from bootlegging. Fascinated by his wealth and devotion, Daisy becomes his mistress. When she kills Myrtle, wife of a garage man and Tom's mistress, in a car accident and runs away, Gatsby comes forward to protect her. He is shot when Tom betrays him falsely to Myrtle's husband out of a gut feeling of jealousy. At his death, all his so-called friends abandon him.

The present study gathered data from the novel, which served as a material source for the qualitative and quantitative analysis of three-word clusters.

Instrument

AntConc 3.4.4w (Windows), one of the most commonly used concordance software, was utilized in the current research to check frequencies and distributions of three-word clusters in *The Great Gatsby*.

Data Collection and Analysis

Data collection in the current research involved the following steps. First and foremost, the current research adopted the corpus-driven approach which is bottom-up and took the data as the center of the analysis. In other words, before looking at the data, the author had no idea or expectations at all about what high-frequency clusters would be extracted from the corpora used for the research. The *N-gram* function in *AntConc 3.4.4w (Windows)* was employed to extract three-word clusters from the corpora. This study concentrated on three-word clusters for that longer clusters are "more phrasal in nature and correspondingly less common" (Biber et al. 1999: 992) though, four-, five- and six-word clusters, that are of more than five occurrences are too few to disclose the textual features in *The Great Gatsby* since the novel consists of only 54876 words. However, three-word clusters that are of 1748 tokens and 213 types occur enough times to be analyzed. Accordingly, the present study eventually selected top 50 three-word clusters of 719 tokens in total to investigate in a manageable way.

Following the collection of high-frequency three-word clusters, they were explored in terms of their frequencies and distributions. To begin with, all instances of clusters were extracted via the retrieval of their concordances, which refer to "a collection of the occurrences of a word-form, each in its own textual environment" (Sinclair 1999: 32). And then the author classified the extracted clusters into the five groups mentioned before based on their functions: Labels, Speech clusters, *As If X* clusters, Body Part clusters and Time and Place clusters. The next step was looking for distributions of individual clusters with the aid of concordance plots. In the wake of investigations above, a great deal of importance was attached to prominent stylistic effects of the novel that clusters exhibit and pragmatic functions of these clusters were discussed.

RESULTS AND DISCUSSION

In response to the research questions put forward previously, the case study below aims to offer an in-depth analysis of three-word clusters and the stylistics effects they exhibit in *The Great Gatsby*, thereby promoting readers' understanding of the masterpiece.

Labels

Clusters in the Labels group "are, or contain, the names of characters, e.g., 'Mr Pickwick and his friends', or in the case of 'The Six Jolly Fellowship Porters', a name for a place: 'a tavern of a dropsical Appearance' (Our Mutual Friend)" (Mahlberg 2007). 2 clusters of this type, the token of which is 18 and accounts for 2.50% of all the clusters obtained, are found in the three-word clusters. Table 1 lists the high-frequency clusters of this group.

As is shown in the Table above, three-word Labels in *The Great Gatsby* takes up a very small percentage among all

the clusters extracted, thus they are not significant enough to disclose certain stylistic effects except for the indication that part of the story is set in New York. This result is in sharp contrast with that of Mahlberg (2007). In his study on five-word key clusters in Dickens's novels, Labels occupy a relatively large percentage, which can be interpreted as Dickens' tendency to characterize people, places and things by labeling them in the stories. On the contrary, Fitzgerald does not prefer to portray things by labeling them. So this group will not be discussed in detail in the present study.

Speech Clusters

Speech clusters contain "a first person or second person pronoun or possessive, which is taken as an indication of interaction" (Mahlberg 2007). Admittedly, there are clusters which are difficult to tell if they are supposed to be in this group in terms of their functions. While coming across such problem, the author went back to the concordance lines to evaluate their pragmatic functions in the context. For instance, all the 14 occurrences of *didn't know* are preceded with *I* or *you* and used in speeches to show interactions of characters according to concordance lines, thus *didn't know*, lacking a first or second pronoun though, is considered as a Speech cluster in this case study. Table 2 lists the 21 clusters, totaling 302 tokens, of this group, which takes up 42.00% of the top 50 clusters.

Conspicuously, a half of the Speech clusters carry negation elements, denoting characters' unwillingness and tendency to reject, successfully helping to create a series of negative characters, such as selfish Daisy, pompous and irritable Tom as well as inactive and confused Wilson. On the other hand, negation-carrying elements build up a negative and melancholy atmosphere, laying the keynote for the

entire novel. Also, the resistance and struggle reflected in the characters' inner world are brought into pointed comparison with prosperous American society, which deepens readers' understanding of the society and interprets the theme of the novel.

Besides, massive occurrences of Speech clusters, taking up 42.00% of the total occurrences extracted, show that, apart from direct narration, dialogues is another crucial way to disclose information. What should be noticed is that four speech clusters *I did not*, *I could not*, *I was not* and *do not think* mainly appear in Nick's narration instead of characters' dialogues, which can be interpreted as interactions between the narrator and readers.

Typically, quantitative as well as qualitative findings of *I do not* and *I did not* that occur most frequently in dialogues are analyzed first below, and then clusters *I want to* and *I could not* that indicate interactions between narrator and readers are demonstrated.

I do not

With the highest frequency among three-word clusters, *I do not* occurs altogether 42 times, encompassing 12 uses by Nick, 8 by Gatsby, 5 by Daisy and Jordan respectively, 4 by Tom, 2 by Wilson and Ms. Wilson respectively and 6 by other characters. As is obviously demonstrated in Table 3, Nick and Gatsby are responsible for 20 occurrences of *I do not*, which occupies nearly a half of the total use.

The concordance plot shows that frequencies of *I do not* are distributed evenly throughout the novel. Usually followed by modality and mental processing verbs like *know*, *want* etc, this cluster plays an essential role in characterizing people, revealing themes and promoting the development of plots through indications of interactions between two characters. Its functions in context are elaborated in following examples.

(1) 'Well, I'm going to tell you something about my life,' he interrupted. 'I don't want you to get a wrong idea of me from all these stories you hear.' (Fitzgerald 2009: 68)

I do not here is uttered by Gatsby asking Nick about his opinion of him on a disconcerting ride. In the past one month, actually, Gatsby doesn't seem to have too much to talk with Nick, which disappoints Nick in whose heart Gatsby's undefined figure fades away gradually. However, driven by extreme vanity and pursuit of materialism, Gatsby eagerly justifies himself by telling Nick his glorified but fabricated past.

(2) 'I don't like mysteries,' I answered. 'And I don't understand why you won't come out frankly and tell me what you want. Why has it all got to come through Miss Baker?'
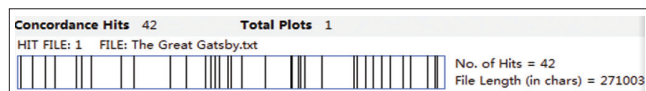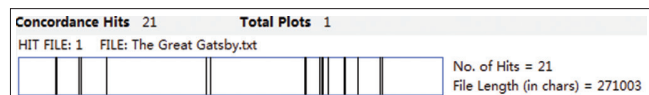
'Oh, it's nothing underhand,' he assured me. 'Miss Baker's a great sportswoman, you know, and she'd never do anything that wasn't all right.'(Fitzgerald 2009: 75)

The above dialogues appear when Nick got to know that Gatsby would request him to do something clandestine for him. Conspicuously, Nick is now getting tired of Gatsby for he seems to obscure certain facts from him. The two successive uses of *I do not* by Nick indicate that he, in an attempt

Table 1. Labels

N	Cluster	F
1	from New York	9
2	in New York	9

(N=No., F=Frequency)

Table 2. Speech clusters

N	Cluster	F	N	Cluster	F
1	I do not	42	12	you do not	12
2	I did not	30	13	I am going	11
3	I want to	21	14	have got to	11
4	I can not	20	15	I was not	11
5	did not know	14	16	I saw that	10
6	I could not	13	17	am going to	10
7	would like to	12	18	do not think	10
8	do not know	12	19	do not you	9
9	I would like	12	20	you did not	9
10	I have got	12	21	you ought to	9
11	was going to	12			

(N=No., F=Frequency)

Table 3. The quantitative findings for *I do not*

Character	Nick	Gatsby	Daisy	Jordan	Tom	Wilson	Ms.Wilson	Others
N	12	8	5	5	4	2	2	4
%	28.57	19.05	11.90	11.90	9.52	4.76	4.76	9.52

(N=No., F=Frequency)

Plot 1. The concordance plot of *I do not*

Concordance Hits 42	Total Plots 1
HIT FILE: 1 FILE: The Great Gatsby.txt	

No. of Hits = 42
File Length (in chars) = 271003

Plot 2. The concordance plot of *I want to*

Concordance Hits 21	Total Plots 1
HIT FILE: 1 FILE: The Great Gatsby.txt	

No. of Hits = 21
File Length (in chars) = 271003

to know the detail about this "mystery", is not satisfied with Gatsby's weird behaviour, which promotes the development of the plot naturally.

Additionally, Nick is actually questioning about the perverse relationships between Gatsby, Miss Baker and himself through such negation. Why should not Gatsby tell Nick directly what to do since it's so convenient? There are even more abnormal relationships hinted in other dialogues, which reflects deeply one theme of the novel: the disruption of American society and decline of American spirit of the day.

I want to

The overall occurrences of cluster *I want to* is 21 times, including 6 uses by Tom, 5 by Gatsby, 3 by Wilson and Ms. Wilson respectively and 4 by other characters. According to Table 4, Tom hits the highest point, followed by Gatsby, and the two rivals in love contribute to over a half of all the occurrences. The natural tendency for *I want to*, as a mental phrase, to "favor" Tom and Gatsby exactly illustrates their struggling inner hearts in various conflicts of the novel.

In view of the concordance plot, cluster *I want to* is distributed mainly in the beginning and latter part throughout the entire novel. It functions as an indication of interaction are exemplified below.

(3)   Wait a minute,' snapped Tom, 'I want to ask Mr. Gatsby one more question.'
      'Go on,' Gatsby said politely.(Fitzgerald 2009: 137)
(4)   'That's a good idea.' I got up. 'Come on, Tom. Nobody wants a drink.'
      'I want to know what Mr. Gatsby has to tell me.'(Fitzgerald 2009: 138)
(5)   Sit down Daisy.' Tom's voice groped unsuccessfully for the paternal note. 'What's been going on? I want to hear all about it.'(Fitzgerald 2009: 138)

The three occurrences of *I want to* in the above excerpts, extracted from the same situation when Gatsby visits Tom and Daisy, are all uttered by Tom. In the first dialogue, Tom actually attempts to express openly that Gatsby is causing rows in his family, so *I want to* here is a euphemistic indication of his provocation. Then Gatsby intends to further infuriate Tom, as a response to which Tom uses *I want to know* to indicate that he is willing to accept such defiance. When Gatsby and Daisy's affair is overt in front of others, Tom is totally shocked by the fact that Daisy never loves him,

he uses "I want to hear about it" to conceal and subside his indignation. Consequently, the successive uses of *I want to* by Tom reveal his emotional changes when confronted with Gatsby, and push the plot towards climax.

(6)   'I don't want to do anything out of the way!' he kept saying. 'I want to see her right next door.'(Fitzgerald 2009: 84)
(7)   'I want to speak to Daisy alone,' he insisted. 'She's all excited now——'(Fitzgerald 2009: 141)
(8)   'I want to wait here till Daisy goes to bed. Good night, old sport.'(Fitzgerald 2009: 155)

The occurrences of cluster *I want to* used by Gatsby are almost all directly related to Daisy as shown in the above examples, manifesting his obsession with her. In the sixth excerpt, Gatsby, expressing his willingness to meet Daisy. The next time Gatsby uses *I want to* is when he is maddened by Tom's words that Daisy doesn't love him, and thus attempts to get reassuring response from Daisy herself to save his face. After Daisy kills Ms. Wilson in a car accident, Gatsby uses *I want to* again to show his firmness, nearly to an extent of madness, to protect Daisy. Frequent uses of *I want to* display a strongly subjective sense of Gatsby and his insane love for Daisy, which is indicative of the novel's tragic ending.

I did not

Nick, as both the narrator and participant of the story, is responsible for 19 out of 30 occurrences of *I did not*, which manifests that he is a discreet person, daring not jump to a conclusion casually. His retrospective narration is limited to first-person point of view, and he can be defined as a reliable narrator as is evident when he claims himself "inclined to reserve all judgments" (Fitzgerald 2009: 1).

According to the concordance plot, *I didn't* are spread across the book, with slightly more occurrences in the latter part. Its functions in revealing Nick's characteristic and enhancing the novel's credibility are exemplified below.

(9)   This was a permanent move, said Daisy over the telephone, but I didn't believe it—I had no sight into Daisy's heart but I felt that Tom would drift on forever seeking, a little wistfully for the dramatic turbulence of some irrecoverable football game.(Fitzgerald 2009: 7)

This part is excerpted from the very beginning of the novel. Although seemingly, it has nothing to do with the hub of the novel for Nick to use *I did not* to show his distrust of Daisy's words, he actually intends to display the rigor and objec-

Table 4. The quantitative findings for *I want to*

Character	Tom	Gatsby	Wilson	Ms. Wilson	Others
N	6	5	3	3	4
%	28.57	23.81	14.29	14.29	19.05

(N=No., F=Frequency)

Plot 3. The concordance plot of *I did not*
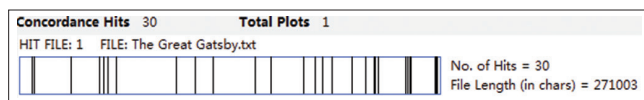
tiveness of his narration even in terms of the slightest thing. In other words, it's owed to his prudent characteristic that he can tell stories objectively to readers and thus acquires their trust.

(10) So I didn't know whether or not Gatsby went to Coney Island or for how many hours he 'glanced into rooms' while his house blazed gaudily on. (Fitzgerald 2009: 88)

(11) For several weeks I didn't see him or hear his voice on the phone—mostly I was in New York, trotting around with Jordan and trying to ingratiate myself with her senile aunt— but finally I went over to his house one Sunday afternoon. (Fitzgerald 2009: 107)

Nick confesses his unawareness of Gatsby's whereabouts by the use of *I didn't know* and *I didn't see or hear*. As can be seen from the above two extracts, Nick, as the narrator, is careful throughout the novel, never telling us things that he could not have known.

I could not

Among the 13 occurrences of *I could not* throughout the novel, 11 are produced by Nick in his narration. From the perspective of its negation-carrying meaning, *I could not* functions similarly to *I did not* in the creation of Nick's characteristic and deepening themes. Nonetheless, this cluster, consisting of a modal verb *could*, demonstrates something more.

Halliday (2000) points out that it's possible to formalize modality to some extent and to establish three values: high, median and low (Thompson 2000: 59). These values are summarized in the following Table:

According to Halliday (2000: 362), *could* is characterized into the low value. And it can be found out that 10 *could not* out of 11 *I could not* in Nick's narration is employed to express his inclination. The choice of inclination of low value (I'm keen, I'm willing, etc.) and the choice of probability of low value (perhaps, may, guess, etc.) are often related to a speaker of mild attitude (Yan 2007: 21). Consequently, Nick uses low value *could not* to indicate his tentativeness, paying respect to the other and leaving room for the other to air his opinion and make up his or her own decision.

As displayed in the concordance plot, cluster *I could not* is distributed mainly in the middle and latter part of the book. Particularly, from the climax of the novel--quarrel between Gatsby and Tom--to the end Gatsby's funeral, *I could not* is frequently employed by Nick, which can be interpreted as his changed emotion towards the story.

At the very beginning of the story, however, Nick pinpoints that Gatsby represents everything for which he has "an unaffected scorn" (Fitzgerald 2009: 2). Gradually, he begins to appreciate, in a reserved way, Gatsby's genuine personality which is in sharp contrast with Daisy's mercenary characteristic. As a result, before the end of the story, Nick shouts to Gatsby that all other people together cannot compare to him. So increasing occurrences of *I could not* in fact demonstrate Nick's inclination to respect Gatsby.

(12) He was clutching at some last hope and I couldn't bear to shake him free. (Fitzgerald 2009: 157)

This is when Nick, in order to prevent Gatsby from being mistaken as the criminal, persuades him to leave the city after Daisy kills Ms. Wilson in the car accident; however, Gatsby, only considering about protecting Daisy, would never agree to do so. Even Nick, as a friend of Gatsby, is moved by so much by his crazy love for Daisy, and can't be hardhearted enough to let him give up. He expresses mildly his admiration for Gatsby for his loyal quality and perseverence in his dreams by using *could not* instead of *can not*, which shows that even he himself is not sure about whether Gatsby's such behaviour, which is destined to be a tragedy, should be advocated or not. Conflicted inner world displayed by *I could not* is actually criticism of the chaotic American society of the day.

(13) I couldn't forgive him or like him, but I saw that what he had done was, to him, entirely justified. It was all very careless and confused. (Fitzgerald 2009: 190)

The above excerpt is nearly at the end of the novel. Nick despises Tom and Daisy for they are indifferent, apathetic and even callous towards everything except money and their own benefits. At the same time, he uses milder *could not* instead of *can not* again to leave certain space for readers to make their own judgments.

Therefore, Speech clusters are employed to indicate interactions between characters as well as narrator and readers. Also, negation-carrying meaning denotes that characters tend to suspect, deny and refuse on the one hand, and build up a melancholy atmosphere, suggesting the failure of American dream on the other.

As If X Clusters

The next group contains one cluster that starts with *as if*: *as if he*.

The phrase *as if*, a half of which preceding the third person pronouns *he, she, it* and *they*, appears 50 times across the novel, illustrating that massive clusters consisting of subjunctive *as if* are employed to portray different characters. Additionally, only 2 out of 50 clusters directed by *as if* are involved in the description of the first and second person pronouns, which demonstrates that clusters preceded with *as if* are usually objective comments upon the appearance, behaviour or mental state of the characters.

Among all the clusters involving *as if*, *as if he* is of the highest frequency, occurring 11 times in total. However, *as if she*, *as if it* and *as if they* appear 6, 3, 2 times respectively, so they will not be discussed in detail. It can be found from the concordance lines of this cluster that the third person pronoun *he* in 9 of its occurrences refers to Gatsby. Fitzgerald reveals Gatsby's distorted disposition through subtle descriptions of his appearance, behaviour and mental state.

According to the concordance plot, cluster *as if he* is spread throughout the whole book. How cluster *as if he* contributes to the portrait of Gatsby is specified in the following examples.

(14) 'Good night.' He smiled—and suddenly there seemed to be a pleasant significance in having been among the last to go, as if he had desired it all the time. 'Good night, old sport.... Good night.' (Fitzgerald 2009: 57)

Gatsby invites Nick to his party and they become acquainted with each other for the first time. Nick doesn't n know much about Gatsby at that time, so he describes him from his own observation to make the character Gatsby vivid. *As if he* here probes into Gatsby's mental state, showing that he in fact organizes this party purposefully and in a mysterious way. It also foreshadows Gatsby's request for Nick's help later.

(15) With his hands still in his coat pockets he stalked by me into the hall, turned sharply as if he were on a wire and disappeared into the living room. (Fitzgerald 2009: 91)

Gatsby is finally going to meet Daisy again under the arrangement of Nick. In the above example, the cluster *as if he* is used to depict Gatsby's behaviour when he is on the point of meeting Daisy. He is so edgy to see Daisy that he is even like being controlled by a wire, downright out of sanity. Such natural reaction exactly reflects the naive and innocent aspect of Gatsby.

Therefore, *As If X* clusters, mainly proceeding the third person pronouns, serve to portray characters through objective comments. Particularly, cluster *as if he* that mainly appears to depict Gatsby reveals his distorted disposition through subtle descriptions of his appearance, behaviour and mental state.

Body Part Clusters

According to Mahlberg (2007), Body Part clusters should consist of at least one part of the human body. However, body language or non-verbal communications can also be covered under Body Part clusters, as outlined by Korte (1997). As listed in Table 7, there are three clusters belonging to this group.

It's notable that the notional verb *look* occurs in 3 of the top 50 retrieved clusters, hitting the highest frequency among all the notional verbs appearing in the clusters retrieved. According to Toolan (2009), the modality and mental processing verbs play a crucial role in narration, and *look* is among the 14 most important such verbs he lists.

With the help of concordance lines, it can be found that the agents of *looked at* include Gatsby, Daisy, Tom, Myrtle and some other minor characters. They are responsible for 8, 7, 3, 3 and 4 occurrences respectively, adding up to

Table 5. Three 'values' of modality (Halliday 2000: 358)

Probability		Usuality	Obligation	Inclination
High	certain	always	required	determined
Median	probable	usually	supposed	keen
Low	possible	sometimes	allowed	willing

Table 6. *As If X* Cluster

N	Cluster	F
1	as if he	11

(N=No.,F=Frequency)

Table 7. Body Part Clusters

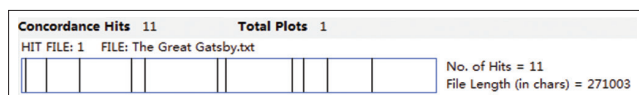| N	Cluster	F
1	looked at me	16
2	he looked at	11
3	she looked at	10

(N=No., F=Frequency)

Plot 4. The concordance plot of *I could not*
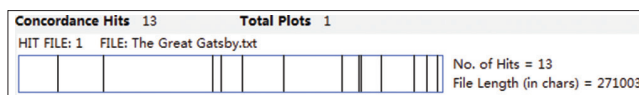
Plot 5. The conco2.116 mmrdance plot of *as if he*

26 occurrences in total, as is shown in Table 8. What has to be specified is that, in order to guarantee the accuracy of frequency, any overlapped occurrences of any two of the three clusters *looked at me*, *he looked at* and *she looked at* is counted as one occurrence only.

Additionally, the frequent occurrences of the phrase *looked at* demonstrate that the narrator Nick, although not talking too much, is adept in revealing personalities of characters by observing others' expressions, which contributes to the objectivity of his narration on the one hand, and exposes his discreet characteristic on the other.

The following are some examples of *looked at* extracted from the novel.

(16) I waited, and sure enough, in a moment she looked at me with an absolute smirk on her lovely face, as if she had asserted her membership in a rather distinguished secret society to which she and Tom belonged. (Fitzgerald 2009: 19)

Nick feels that Daisy is not communicating with him sincerely, and the way she looks at him further confirms his guess. Here the action *looked at* betrays Daisy's mental state: she advertises herself as nobility belonging to the upper-class society. *Smirk* shows her feeling of being privileged over Nick, which contradicts what she says to Nick before, reflecting her materialist value and artificial personality.

(17) He looked at me sideways—and I knew why Jordan Baker had believed he was lying. (Fitzgerald 2009: 68)

When Gatsby tells Nick about his life experience, trying

Table 8. The quantitative findings for agents of *looked at me, he looked at & she looked at*

Characters	Gatsby	Daisy	Tom	Myrtle	Others
N	9	7	3	3	4
%	34.62	26.92	11.54	11.54	15.38

(N=No., F=Frequency)

Table 9. Time and Place clusters

N	Cluster	F	N	Cluster	F
1	for a moment	33	4	for a while	13
2	half an hour	16	5	into the room	11
3	after a moment	15	6	a moment I	10

(N=No., F=Frequency)

to break away from those "scandals" about him out of his vanity, Nick notices that Gatsby squints at him. This seemingly trivial behaviour denotes that Gatsby is actually guilty inwards for he is fabricating something not existing. To keep his narration reliable and objective, Nick discloses Gatsby's guilty mantle state and weird characteristic through his observation of Gatsby's behaviour.

Hence, Body Part clusters are mainly concerned with the phrase *look at*, which demonstrates that Nick tend to reveal characters' mental process by observing others' expressions, which contributes to the objectiveness of the novel. This tendency exposes his discreet characteristic on the other hand.

Time and Place Clusters

Time and Place Clusters are clusters referring to time or place. When clusters contain features of both Labels and Time and Place Clusters, the Labels category overrules Time and Place; an example is 'from New York' (in Table 1 above), which contains a particular name of place, is thus classified as Label.

There are 87 occurrences of Time and Place clusters in total. As Table 9 shows, 5 out of 6 clusters belonging to this group are Time clusters, occupying 77.55%, which definitely accounts for the smooth motion of narration. This result is also in accordance with the fact that *The Great Gatsby* is developed in chronological order. From qualitative perspective, all the five Time clusters, namely *for a moment*, *half an hour*, *after a moment*, *for a while* and *a moment I*, denote relatively short time span, and there are even three clusters including the word *moment*.

With the aid of massive Time clusters characterized by indication of short time span, the progressive layers of plots tightly interlocked are developed in a compact way and full of suspense, which makes the whole story, although consisting of only nine chapters, extremely compelling. However, due to the frequent occurrences of this group in novel narration, it will not be discussed from qualitative perspective in detail here.

In addition to these five groups, there are fifteen clusters that do not fit into either of the groups. But like Mahlberg (2007), this study will not provide more detailed classifications.

In the light of the preceding discussion, it becomes clear that three-word clusters play an important role in exhibiting stylistic effects of *The Great Gatsby*,like shaping the characterization of Gatsby, Daisy, Tom, Nick, to name a few, and deepening the themes as well as accounting for the plot development in *The Great Gatsby.*

CONCLUSION

The present study provides a detailed description of high-frequency three-word clusters in Fitzgerald's *The Great Gatsby*. The use of the corpus methodology in the research has made some uncovered valuable stylistic information hidden in the novel. The top 50 three-word clusters can be classified into Labels, Speech clusters, *As If X* clusters, Body Part clusters and Time and Place cluster. The most typical ones include *I do not*, *I want to*, *I did not*, *I could not*, *for a moment*, *looked at me*, *as if* he and so on. Certain characters and plots in *The Great Gatsby* are found to yield high frequencies of these clusters, which are unevenly distributed throughout the novel. Many of those clusters bear common textual functions, which contributes more or less to the characterization, themes, and plot development in *The Great Gatsby* as well as the objectiveness of the novel.

ACKNOWLEDGMENTS

REFERENCES

Allan, R. (2016). Lexical bundles in graded readers: to what extent does language restriction affect lexical patterning?. *System,* 59, 61-72.

Biber, D., Johansson, S., Leech, G. et al. (1999). *The Longman Grammar of Spoken and Written English*. London: Longman.

Cao, Y. (2017). *A Corpus Stylistic Study of Theodore Dreiser's Novels (*Unpublished master thesis*)*. Tianjin University of Science Technology, Tianjin.

Fischerstarcke, B. (2009). Keywords and frequent phrases of Jane Austen's pride and prejudice A corpus-stylistic analysis. *International Journal of Corpus Linguistics*, 14(4), 492-523.

Fitzgerald, F. S. (2011). *The Great Gatsby*. London: Penguin Books Ltd.

Grabowski, K. (2015). Keywords and lexical bundles within English pharmaceutical discourse: A corpus-driven description . *English for Specific Purposes*, 38, 23-33.

Halliday, M. A. K. (2000). *An Introduction to Functional Grammar*. Beijing: Foreign Language Teaching and Research Press.

Hardy, D. (2003). *Narrating Knowledge in Flannery O'Connor's Fiction*. California: University of South California Press.

Huang, K. (2015). More does not mean better: Frequency and accuracy analysis of lexical bundles in Chinese EFL learners' essay writing. *System*. 53(9), 13-23.

Korte, B. (1997). *Body Language in Literature*. Toronto: University of Toronto Press.

Mahlberg, M. (2007). Clusters, key clusters and local textual functions in Dickens. *Corpora*, *2*(1), 1-31.

Mahlberg, M. (2010). Corpus linguistics and the study of nineteenth century fiction. *Journal of Victorian Culture*. 15(2), 292 -298.

Meunier, F. & Gouverneur, C. (2007). The treatment of phraseology in ELT textbooks. *Language & Computers*. 61, 119-139.

Nekrasova, T. M. (2009). English L1 and L2 speakers' knowledge of lexical bundles. *Language Learning*. 59(3), 647–686.

Qin, J. (2014). Use of formulaic bundles by non-native English graduate writers and published authors in applied linguistics. *System*. 42(1), 220-231.

Scott, M. (2004). *WordSmith Tools. Version 4.0. Manual*. Oxford: Oxford University Press.

Sinclair, J. (1999). *Corpus, Concordance, Collocation*. Shanghai: Shanghai Foreign Language Education Press.

Anna, S. C., Kathy, C. & Walter, V. H. (2011). Seeing a phrase "time and again" matters: the role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology, Learning, Memory and Cognition*. 37(3), 776-784.

Stubbs, M. (2005). Conrad in the computer: examples of quantitative stylistic methods. *Language Literature, 14*(1), 5-24.

Tabata, T. (2002). Investigating stylistic variation in Dickens through correspondence analysis of word-class distribution. *Language & Computers*. 165-182.

Toolan, M. J. (2009). *Narrative Progression in the Short Story: A Corpus Stylistic Approach*. Amsterdam: John Benjamins Publishing Company.

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: evidence from self-paced reading and sentence recall tasks. *Language Learning*. 61(2), 569–613.

Wang, Y. (2017). Lexical bundles in spoken academic elf. *International Journal of Corpus Linguistics*. *2*2(2), 187-211

Wray, A. (2002). *Formulaic Sequences and the Lexicon*. Cambridge: Cambridge University Press.

Yan, S. Z. (2007). *An Analysis of interpersonal meaning in Wuthering Heights from the perspective of functional grammar* (Unpublished master thesis). Xijiang University, Xinjiang.

Zuo, J. M. (2006). *A Brief Introduction to American Literature*. Qingdao: China Ocean University Press.