
| RESEARCH ARTICLE

A Comparative Study of Thematic Progression in Human-written and ChatGPT-generated Texts

Xiulin Gao

College of Foreign Studies, Jinan University, China

Corresponding Author: Xiulin Gao, **E-mail:** rhgao199214@163.com

| ABSTRACT

The wide use of ChatGPT has raised concerns about accuracy of the essays it produced. This paper compared the thematic choices and thematic progression patterns in human-written and ChatGPT-generated texts based on the systemic functional grammar and analyzed their coherence and discourse structure. The finding revealed that ChatGPT uses more repeated marker Themes in topical Theme making complement as Theme, but less interpersonal Themes than human which indicates its lack of interaction with readers and subjectivity of expressing its opinions. It also uses less various types of textual Themes than human who employs richer ways to extend, combine and conclude the clauses. Thematic progressions are much more frequent in human-written essays and linear pattern accounts for the largest percent while constant pattern is more frequent in ChatGPT. Human can adopt different thematic progressions and other thematic structures to make a coherent text but ChatGPT mostly repeat the same type of clause and even produce redundant information or uses inaccurate thematic progression making the text a list of information.

| KEYWORDS

Thematic progression; ChatGPT-generated texts; Human-written texts.

| ARTICLE INFORMATION

ACCEPTED: 01 July 2026

PUBLISHED: 15 August 2026

DOI: 10.32996/ijllt.2026.9.8.13

1. Introduction

Theme and rheme are two concepts of the actual division of the sentence which was first put forward by Mathesius. Later Halliday developed them further and presented the textual function, explaining the organization and structure of discourse by theme and rheme. As the point of departure of sentence, theme starts the message and is the first constituent of the sentence. Halliday (1985) originally suggested that Theme is "what the clause is about". Theme is subject, base, the topic of the sentence. Rheme, on the contrary, is the end of the sentence which is the rest part of the sentence and describes, specifies the content of Theme. The subject always tends to be seen as the theme because it often appears at the beginning of the sentence. It is normal to regard the subject as theme so it is called unmarked Theme but other constituents like adjunct, complement are placed in the position of Theme because of contextual reasons which are said to be marked Theme. For instance:

- (1)a, Mary has been told not to touch the plant which is poisonous.
b, Last night, a man robbed the bank.

The Theme in 1a is "Mary" which is unmarked Theme while "last night" is a marked Theme. Halliday (1985) divided the Theme into two types: simple Theme and multiple Theme in *An Introduction to Functional Grammar*. Simple Theme can't be divided further and does not have internal structure. It is seen as a whole which is usually expressed by nominal group, adverbial group or preposition phrase. Different from simple Theme which only represent one of the metafunction, experiential which is also called topic theme, interpersonal or textual function, multiple Theme can include more metafunctions. The experiential theme is mostly a participant, process or circumstance. Interpersonal and textual Theme precede it. If three metafunctions are included,

Copyright: © 2026 the Author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) 4.0 license (<https://creativecommons.org/licenses/by/4.0/>). Published by Al-Kindi Centre for Research and Development, London, United Kingdom.

the order is: textual Theme, interpersonal Theme and experiential Theme. The textual Themes include conjunctive adjuncts, conjunctions and continuatives (Halliday & Matthiessen, 2014: 107). And the interpersonal Themes that express writer's attitude and judgement of the message. They include vocative, modal adjuncts and finite verbal operators. They have four types of modal orientation: "explicit subjective", "explicit objective", "implicit subjective", and "explicit-implicit subjective". "Explicit subjective" interpersonal Themes are projecting mental clauses such as "I think...". "Explicit objective" interpersonal Themes use rational clauses such as "It is likely that...". "Implicit subjective" interpersonal Themes have modal adjunct such as probably or modal verb like should. "Explicit-implicit subjective" interpersonal Themes include preposition phrases that suggest subjectivity such as "in my opinion".

- (2)a, That fat boy has eaten the whole birthday cake.
b, But, actually, no one has got injured.

In 2a, the simple Theme "that fat boy" only expresses the experiential function. In 2b, the multiple Theme include the textual element "but", the interpersonal element "actually" and the experiential element "no one".

Except for the theme, the orders of Theme and Rheme vary from different registers. thematic progression is a term used to indicate the different construction of Theme and Rheme which is helpful when analyzing the coherence and cohesion of text. There are three main types of progression. In constant progression, the Theme of one clause refers back to the former Theme. In linear Theme, the Theme relates back to one or more constituent in the Rheme of the proceeding clause. In derived progression, Theme is related back to "a "hyper-Theme" which establishes the topic for a longer stretch of text" (Thompson, 2014:175). Other progressions are also used frequently. In centralized progression, the Themes are different while the Rhemes are the same. In intercrossing progression, the Rheme repeats the former Theme in the proceeding clause. Studying the different types of thematic progression helps us have a deeper understanding of the internal logical structure of discourse and write a more coherent text.

Coherence, a typical and basic characteristic of discourse, is in the mind of the writer and reader, which is vitally significant for writing and for reader to know what the author wants to express. It is a semantic concept which can be divided into two types: within and out of language. Inside the language, it manifests itself in different levels of interwoven and interconnected semantic relations. Such semantic relations are manifested in structural relations, organic relations between constituents in discourse and the componential relations. Out of language, it is the relation between the discourse and the situation. Therefore, to make a discourse coherent, there several conditions: clauses should be interconnected and reflected with each other, constituents should be connected to make discourse an organic whole, and the discourse should correlate with the external situation. This paper will discuss the coherence from the perspective of thematic progression.

ChatGPT, an artificial intelligence language tool released to the public in November of 2022, has been widely used by people all over the world. It can give quick and detailed response within a few seconds and has been used by many students to finish their homework and become a guide for writing. But many questions have been raised about its logic and accuracy.

2. Literature Review

With the development of large language model, many scholars have been focused on the application of AI and combined it with other subjects such as linguistic, pedagogy and others. David Baidoo-Anu and Leticia Owusu Ansah (2023) revealed the benefits and limitation of AI in teaching and learning and argued that ChatGPT and other generative AI will continue revolutionizing the current education system. Robert Godwin-Jones (2022) studied the use of AI in writing and learning and suggested that learners should pay attention to the learning environments, context and learner agency and equity. Moreover, various scholars compared the AI-generated texts and human-written articles. Giordano Tudino, Yan Qin (2024) found that in ChatGPT's handling of academic discourse, there are overuse of infrequent vocabulary, limited use of subordination and syntactic and semantic homogeneity. Heather Desaire, Aleesa E. Chua, Madeline Isom, Romana Jarosova, David Hua (2023) reported a method for discriminating text generated by ChatGPT from (human) academic scientists, relying on prevalent and accessible supervised classification methods and claimed that scientists write long paragraphs and have a penchant for equivocal language, frequently using words like "but," "however," and "although." Tony Berber Sardinha (2023) suggested that AI's ability to capture the intricate patterns of natural language remains limited. Ismail Dergaa, Karim Chamari, Piotr Zmijewski, Helmi Ben Saad (2023) highlights the need for comprehensive discussions on the potential use, threats, and limitations of these tools, emphasizing the importance of ethical and academic principles. Yang Linxiu and Ping Jiapeng (2024) argued that the abstract produced by AI neglects the boosters and attitude markers which introduces semantics vagueness and thus undermines the authority and validity of the academic writing.

Few scholars studied the AI-generated texts from the perspectives of textual metafunction. Shu Yang, Shukun Chen, Hailin Zhu, Jiayi Lin, Xi Wang (2024)'s analysis of the textual Theme reveals that the machine tends to use clause-initial conjunctive adjuncts of addition to extend the clause to the previous text and the concessive signals to reiterate information that has been stated earlier and employ fewer condition signals as the textual Theme and it seldom uses modal adjuncts.

For connecting the thematic progression with coherence, there already have numerous studies. Zhang Delu and Liu Hongming (1994) suggested that the choice of theme is one of the important conditions of discourse coherence which should make clauses interconnected. Zhu Yongsheng (1995) found that different types of thematic progression have different roles in making discourse coherent. Ma Jin (2001) proposed that the thematic progression benefits students and teachers from teaching and learning coherent writing.

With the advanced improvement of ChatGPT, people are supposed to make better use of it and take its advantages and try to delete its disadvantages. This study compares ChatGPT-produced texts with human-written articles and analyzes the accuracy and coherence of them with the thematic theory and try to answer the questions: what are the differences between AI-generated and human-written texts in choosing Theme and thematic progression; It is the AI-generated text coherent?

3. Methodology

3.1 Research Questions

This study aims to compare the thematic choices and thematic progression patterns in human-written and ChatGPT-generated texts, comparing the differences between human writing and AI writing. In specific, the study tries to answer the following questions:

- (1) What are the differences of thematic choices in human-writing and ChatGPT-generated essays?
- (2) How do thematic progression patterns differ in human-authored essays and ChatGPT-generated essays?

3.2 Data Collection

In this study, two corpora are established. One is the human-written articles written by professional authors in the large free website-----studymoose (<https://studymoose.com>) which provides numerous free essay example in any topic. The authors in this website are all verified experts in many fields and have been examined and tested. 10 essays are collected ranging from education to society with around 400 words every essays. Some of the writers express their opinions about some problem in school or society while others may give their own suggestions in some aspects. To avoid the influence of ChatGPT and ensure the validity of the statistics, the essays are all written and published before the release of ChatGPT, that is before 2022.

The AI tool that was used is ChatGPT5.2. It has been improved and provides more accurate answers and has less bugs than the former versions of ChatGPT. It is a large multimodal model which can exhibit human-level performance and give text output. 10 questions are raised according to the topic of the 10 essays selected in studymoose and ChatGPT's instant answer are collected. The prompt we used is: "Please write a [number]-word essay based on the topic of [topic]." A new chat was created for each essay generation task to avoid the possible influence of the previous order and information. Each produced essay is about 300-400 words, which is similar to the human-written essays. Here is some basic data about the two corpora.

Table1: Basic Information about the Data

Basic Information about the Data	Human	ChatGPT
Total number of words	4,525	4,110
Average number of words per essay	452	411

3.3 Data Analysis

The study based on the two corpora uses UAM CORPUS TOOL6 to signalize the Theme and Rheme, analyze the quantity of them and compare the exact words used as Theme or Rheme and subtypes of Themes in two corpora.

The qualitative and quantitative methods are adopted. The quantitative approach is used to count the number of different types of Themes and thematic progression and compared them in the AI-generated and human-written essays, analyzing the frequency. The qualitative method is used for explaining the thematic choices of ChatGPT and human, how they adopt thematic progression patterns and comparing the coherence of their texts.

4. Comparative Analysis and Results

4.1 Theme Analysis

In this section, thematic choices of both human and ChatGPT will be analyzed and compared together. The topical, interpersonal and textual Themes will be respectively explained.

Table2: Different Types of Themes in the Essays

Theme Type	Human		ChatGPT	
Topical Theme	295	(74.87%)	281	(87.81%)
Textual Theme	77	(19.54%)	36	(11.25%)
Interpersonal Theme	22	(5.58%)	3	(0.94%)
Total	394	(100%)	320	(100%)

From the above two charts, we can see that the total number of words in human-written texts is larger than that of ChatGPT so the Themes that human uses are much more than the ChatGPT uses. The AI tends to produce short and brief answers. But the proportions of the topical, textual and interpersonal Themes in them is similar to each other. The most frequently used Theme in both human-written and ChatGPT-generated texts is topical Themes which express the main and important information in the discourse. Textual Theme which is second large number of Theme choices connect clauses to make the discourse coherent so both human and ChatGPT choose them. The last one is interpersonal Themes that are less used to show the author's attitudes. However, human uses more textual and interpersonal Themes but the robot uses more topical Themes. These may indicate that the AI emphasizes more on the content and has less involvement while human pays more attention on the structure of discourse and the author's attitude. The robot does not put emphasis on the interaction between the writer and the reader and may less coherent than the human discourse. More evidences and details will be showed below. The textual, interpersonal and topical Themes will be respectively compared.

4.1.1 Topical Theme Analysis

Topical Themes indicate the experiential content of the clause, and signal the actual starting point of the message (Thompson, 2014). They can be unmarked or marked. They help to signal "the maintenance or progression of 'what the text is about' at that point" (Thompson, 2014: 171).

Table3: Types of Topical Themes in the Essays.

Topical Themes	Human	Frequency	ChatGPT	Frequency
Unmarked	320	81.43%	222	69.38%
Marked	73	18.58%	98	30.63%
Total	394	100%	320	100%

From table3, we can tell that the unmarked Themes and marked Themes in human-written essays account for 81.43% and 18.58% and the robot's unmarked Themes are less but the marked Themes are more. They both have more unmarked Themes since they are the most typical to serve as Themes. The pronouns such as it, I and they are truly common subject as Themes.

(3) a, Students can take classes whatever time they want. (Text 1 of human-written essays)

b, Students can often choose their classes and plan their day, but this flexibility requires careful planning to balance academics, work, and personal life. (Text 1 of ChatGPT-generated essays.)

For instance, "students" is served as the Themes to indicate the object of the action as the starting point of a clause.

The machine writing may use more other constituents as Themes rather than the content to highlight other constituents. Marked themes either "express some kind of setting for the clause or express a feature of contrast" (Halliday, 1994: 48). Adjunct and complement are always treated as the unusual Themes as the marked Themes. Here are the exact statistics of adjunct Themes and complement Themes:

Table4: Adjunct and Complement as Themes in the Essays.

Types of Marked Themes	Human		ChatGPT	
Adjunct as Themes	52	71.23%	58	59.18%
Complement as Themes	21	28.77%	40	40.81%
Total	73	100%	98	100%

Human and AI both tend to use more adjunct as Themes than the complements. It is common to put the location, time or other special conditions on the beginning of a clause to fulfill the contextual needs. Complements are less significant compared with other constituents such as subject. Adverbials are both used by human and AI as Themes.

- (4) a. In high school, there is a set time in which school is in session every day in school. (Text 1 of human-written essays)
 b. In high school, students often have a structured schedule set by teachers and parents. (Text 1 of ChatGPT-generated essays.)

For example, “in high school” is adverbial about the location, used by both human and the robot to address the importance of place. According to the corpus, human writes more if-clauses as Themes and the ChatGPT uses more when-clauses and while-clauses as Themes.

- (5) a. If banning phones in classrooms can save even one teenager’s life, then banning phones is far worth the cost. (Text 9 of human-written essays)
 b. While smartphones offer convenience and access to information, their presence in classrooms often does more harm than good. (Text 9 of ChatGPT-generated essays.)
 c. When students are glued to their screens, they miss opportunities to interact face-to-face with peers and teachers. (Text 9 of ChatGPT-generated essays.)

In4a, writer had the if-clause and the gerund as the subject. ChatGPT had the simple when and while clause as the Themes. Human is more able to use various adjuncts as Themes. Beside the if-clause, there are many preposition groups such as by doing and for example. The adjunct Themes that the robot uses are much more longer sentences which may not as brief and clear as those written by human.

4.1.2 Interpersonal Theme Analysis

Interpersonal Themes are used for indicating the author’s attitude towards things. They signal the writer’s awareness of the interaction with the reader and function as interactive thematic threads running through the text (Thompson, 2007: 690). In table2, the interpersonal Themes account for 5.58% and 0.94% in human-written and ChatGPT-generated essays. Both human and robot focus on expressing ideas and opinions but robot has much less awareness of this.

Table5: Words of Interpersonal Themes.

Rank	Words	Human		Robot	
1	is	11	50%	1	33.33%
2	are	2	9.09%	0	0%
3	’s	1	4.55%	2	66.67%
4	maybe	2	9.09%	0	0%
5	generally	1	4.55%	0	0%
6	often	1	4.55%	0	0%
7	usually	1	4.55%	0	0%
8	comfort	1	4.55%	0	0%
9	personally	1	4.55%	0	0%

Table6 presents the use of interpersonal Themes of human writers and ChatGPT. As it shows, the most frequently used Themes is the “explicit objective” realization such as “It’s essential to...” which increases the validity and objectivity of the texts. Human

writers also use adverbials to signal the frequency and probability which adds the accuracy and the authenticity of the essays. The robot writing is just lack of these words that makes its essays not as real and vivid as those of human.

(6)a, Personally, I find solitude more inspiring and peaceful although I am a social person and have a lot of friends. (Text 3 of human-written essays)

b, I personally could benefit from practicing more confidence and self-assurance. (Text 10 of human-written essays)

In (6), the use of “explicit subjective” situation indicates that human writers have the writer awareness to address their ideas towards things and the interaction between writer and reader but ChatGPT has no such awareness.

4.1.3 Textual Theme Analysis

Textual Themes include conjunctions and conjunctive Adjuncts, linking clauses together to make the discourse logically and semantically coherent. Through comparing the textual Themes of human-written and ChatGPT-generated essays, we can explore the coherence and cohesion of the texts.

Table6: Major Textual Themes Types in the Essays.

Textual Themes	Human		ChatGPT	
Addition	12	15.58%	5	13.89%
Appositive	5	6.49%	1	2.78%
Consequence	8	10.39%	1	2.78%
Adversative	9	11.69%	4	11.1%
Concluding	7	9.09%	0	0%
Sequencing	5	6.49%	0	0%

In table2 and table6, the ChatGPT has less textual Themes than human. The most frequently used subtypes of textual Themes are addition and the second one is adversative. Human has more kinds of subtypes textual types than the robot uses. There are other types of textual Themes which are not showed in the chart. Table7 presents the mostly used words of textual Themes:

Table7: Words of Textual Themes in the Essays.

Human				ChatGPT			
Rank	Words	Frequency		Rank	Words	Frequency	
1	If	11	14.29%	1	While	11	30.56%
2	so	7	9.09%	2	when	6	16.67%
3	and	5	6.49%	3	whether	6	16.67%
4	when	5	6.49%	4	And	3	8.33%
5	For example	4	5.19%	5	however	3	8.33%

Table7 shows that human uses more conditional clauses to propose a condition that may exist and present more potential situations. Robot tends to use more words like while and when which is commonly used but too normal. ChatGPT frequently uses “while” to compare two opposite conditions and “when” to address some different conditions. Thus, ChatGPT extends the information and compared different situations. In order to make the texts more logical, human supposes distinct situations with conditional clauses and has more arguing and structuring words such as conclusion and apposition.

(7)a, Another difference is the teachers in high school and professors in college..... Instead, it will be the student's responsibility to get notes from a classmate. In addition, if students want one on one time with their professor, they would not be able to stay

after class like in high school: they will have to make an appointment during the professor's office hours. (Text 1 of human-written essays)

b, While high school offers extracurricular activities, college takes it to the next level. From internships and research opportunities to clubs and study-abroad programs, the possibilities are endless. (Text 1 of ChatGPT-generated essays.)

In 7a, human applies more conjunctive adjunct to connect clauses in the clause complex. However, ChatGPT just lists "while" or even uses nothing to link clauses. Therefore, the human essays are more logical and cohesive than the ChatGPT essays.

4.2 Thematic Progression and Coherence Analysis

The different choices of Theme and Rheme show the semantic connection and logical relation between sentences, the information structure and how the writer constructs the discourse structure to make a coherent text. Constant patterns create cohesion by reiterating the same element as the Theme; linear patterns achieve cohesion in the text by "building on newly introduced information" (Eggins, 2004: 325).

Tabel8 presents the thematic progression patterns of human-written and ChatGPT-generated texts. We can tell that human uses more thematic progression than ChatGPT so it can be inferred that human is much more aware of the choices of Theme and pay attention to the coherence of the text. In human-written essays, the most frequently used pattern is linear progression whose percentage is higher than that of ChatGPT's generated essays. The second frequent pattern is constant progression but its percentage is lower than ChatGPT's constant progression. The third one is derived progression which is applied more frequently than ChatGPT's derived progression. In ChatGPT-generated texts, the most frequent pattern is linear progression and the second frequent is constant progression. These two pattern have almost the same frequency percent. And the last frequent pattern is derived progression.

Table8: Thematic Progression of Human-written and ChatGPT-generated Essays

Type	Human		ChatGPT	
	Quantity	Frequency	Quantity	Frequency
Constant	25	35.21%	17	43.59%
Linear	37	52.11%	18	46.15%
Derived	9	12.68%	4	10.26%
Total	71		39	

ChatGPT uses the constant pattern more often than human. The constant pattern creates a "static" style of writing (Enkvist, 1973), making the text read as incoherent as a list (Ebrahimi & Khedri, 2011). The repetition of the same Theme will make the essay tedious and boring. The essays that Chat GPT created are more like a list of some topics rather than a coherent passage. It just collects the information from the big data on the Internet and does not consider the thematic choice, use of thematic progression or the information structure as the human does.

(8)a, Positive self-talk—encouraging and affirming thoughts directed toward oneself—can be a game-changer in achieving personal growth, overcoming challenges, and maintaining mental well-being. Let's explore the benefits of positive self-talk and practical techniques to cultivate it.

Positive self-talk helps build self-esteem by reinforcing a belief in your abilities.

.....

Positive self-talk fosters a proactive attitude that leads to goal achievement.

Consistent positive self-talk contributes to a more optimistic outlook on life.

(Text2 of ChatGPT-generated essays)

b, How often in a day do you find yourself talking to yourself mentally? Me, so very often and in a positive way! And this is not surprising since research has already shown that talking to yourself positively can improve your self-esteem, management of stress, wellbeing.

Self-talk is also known to cut any indications of personality disorders, anxiety, and depression, besides improving body image, as well as treating people with eating complaints. Then again, if you mostly think negatively on yourself, the results are the opposite. (Text2 of human-written essays)

In (8)a, “self-talk” has been the Themes for several times at the beginning of the clause which stops the discourse from developing and becoming varied and rich. This might show that AI uses less various words than human writers and its ability of applying the thematic progression pattern and constructing the information structure is limited. Human, on the other hand, pays attention to how to arrange the information and guides the readers to understand the articles better and faster. Human writers are able to use different words as Themes and thematic progression flexibly and there are more conjunctions in human-written essays to lead the readers to know the author’s opinions gradually and deeply. For instance, in (8)b, linear progression is used at the beginning of the essay and there are no repeated words as Theme. Textual Themes like “and” are used for showing the logical relation between clauses such as parataxis. Although the topic is self-talk, human writer won’t always repeat it as ChatGPT does but presents it in other ways like “talking to yourself” in (8) b. Thus, human-written essays are more dynamic while AI-generated essays are more static.

ChatGPT only has single and simple method to use thematic progression patterns. Human writers are able to have the thematic progression in various ways. With the analysis of the thematic progression pattern, we can know the distribution of the known information and new information and the interaction between them (Zhu Yongsheng,1995). Human usually follows the natural reading habit which is reading from the old information to the new one. So linear progression is often employed in writing.

(9) a. Research suggests that people form judgments within seconds of meeting someone. These initial perceptions are often based on appearance, body language, and communication style. (Text3 of ChatGPT-generated essays)

b, It’s a natural occurrence as a human being to make unintended judgments based on appearance. This includes what you wear, how appropriately you dress, and how you carry yourself. The way you present your personal appearance can be a disadvantage or it can be an asset. (Text3 of human-written essays)

c, You can learn their culture by interacting with them, tasting their food, visiting their stores, restaurants, temples and museums. From the museums that display the history of the local city, you can learn more and also break down the stereotypes and misperceptions from the books. (Text6 of human-written essays)

d, When we are alone, most of our day comes back to our mind like a movie and it is the best time to evaluate everything. This way we consciously reinforce the good parts and try to change the bad ones. For example, if we have a fight or a strong argument during the day, we can recall the incident and try to judge the whole situation objectively without strong emotions. (Text10 of human-written essays)

Most linear progression patterns in AI-generated essays have demonstrative pronoun and personal pronoun as the Theme to refer to the Rheme in the proceeding clause. In (9)a, “these initial perceptions” refers to the “judgement”. Like the robot, human also uses pronoun to avoid repeating the same things but human writers employ more various ways to indicate information mentioned before. Synonymy, hyponymy and repetition are frequent in linear pattern of human-written essays. For example, “this” relates to “appearance” and “the way you present your personal appearance” refers to the content before in (9) b. “Museums” is both Rheme in the first clause and Theme in the second clause in (9) c. In (9)d, “this way” signal the action in the previous clause and “have a fight or strong argument” is a more concrete example corresponding to the Rheme “bad ones”. The writer first showed the situation of being alone, then the result of it, and finally made an example to explain the result. The thematic progression structure of these three instances is “R1=T2, R2=T3, R3=T4.....”. Theme is related to the Rheme of the previous clause. The writer initially claimed his ideas and then explained them in details. The information was arranged from old to new and abstract to concrete which conforms to people’s cognitive habits and makes reading more convenient. Such information structure constructs an information chain in which new information is becoming a known information that becomes a starting point introducing another new information (Zhao Pu,1998). ChatGPT-generated essays, on the other hand, are far less organized and coherent. Unlike the human writers, ChatGPT mostly uses only one thematic progression pattern in a paragraph and has less explanation and details to verify the argument.

(10)a, Good leaders inspire trust not by making grand speeches but by leading through example. A leader who works hard, listens to their team, and follows through on commitments earns respect through their actions. (Text8 of ChatGPT-generated essays)

For instance, in (10)a, there is only constant progression and only two clauses in this example. In addition, the second clause shares the similar meaning of the first clause which seems to be redundant. To claim the argument that action speaks louder than words, ChatGPT listed many examples in which the leadership above is one of them. Aspects of the information are separated and are independent. These essays have very less complete thematic progression constructions but single thematic progression pattern. Therefore, AI-generated texts are not coherent semantically.

The Themes chose in a coherent discourse are beneficial for the interconnection between different parts in the discourse. Different levels have distinctive Themes which attached to and serve for the Themes or Rhemes in the last level and all levels belong to the topic of the whole discourse (Zhang Delu& Liu Hongmin, 1994).

Derived progression has the function of building basic framework of the discourse and making every level serves for not only the main topic but also Theme or Rheme at its last level. There is hyper Theme in an introductory sentence that sets up the frame for a sequence of following sentences. For example, in text4 of human-written essay, the last sentence in the first paragraph is "It is possible to find happiness in the following places; family, health, and good education," which lead to the following points of the cause of happiness, the topic of the essay. Therefore, the following Themes in next levels are: "Firstly, family"," Secondly, health" and "Finally, ...", which respectively explained why family, health and education can bring happiness. Every Themes in second level refers to the Rheme in the first level, making the discourse coherent and has a clear structure. Also, it implies that the writer has a clear sense of how the text will unfold different points and aims to guide the reader cooperatively through the text (Thompson, 2014). ChatGPT, however, doesn't have such awareness and uses incomplete derived progression. Take text2 as an example, when stating the techniques of cultivate positive self-talk, ChatGPT initially indicated that "the first step is awareness" in the first aspect of identifying negative patterns without introductory sentence or "the second, third step". The Themes in the second level are independent rather than corresponding to the Themes or Rhemes in the last level. Coherence can be presented as semantic relations of different levels interrelate with each other. (Zhang Delu& Liu Hongmin, 1994). Since ChatGPT pay no attention to such interrelationship, AI-generated texts doesn't have clear and coherent discourse structure.

Besides, there are other special type of thematic progression pattern in human-written essays. For example, intersectant pattern in which Theme in the proceeding clause is the same as the Rheme in the next clause and same Rheme pattern in which two clauses share the same Rheme.

(11)a, People who keep going out with friends all the time will suffer from health problems and spiritual issues. When I say health problems, I mean eye and ear irritation, weakness in the body, and problems in the nervous system. Moreover, they will feel a lot of emptiness in their souls because they are used to being with other people and get scared of being alone. (Text10 of human-written essays)

b, Education is one way of getting that dream job and education is an essential cog in the wheel to living comfortably and happily. One English survey that included over 15,000 participants revealed that 81 percent of people who had achieved a good level of education had a high level of life satisfaction. (Text5 of human-written essays)

In (11)a, "health problems" is the Rheme of both the first and the second clause. Additionally, the Theme in the fourth clause "they" refers to the Theme "people" in the first clause. In (11)b, firstly the constant progression pattern is used and the first two Themes "education" appeared in the third clause as Rheme. This again clarifies that human can combine different thematic progression patterns to connect different clauses like a chain. ChatGPT can only repeat the same pattern and the same sentence types.

Also, human makes use of some special thematic structures. Thompson (2014) declared several types of it: thematic equatives, predicated Theme, thematized comment and so on. Take thematic comment as an example, Thompson (2014) thought if one only simply labels "It" as Theme it will obscure the method of development of the text. "It is very true that..." in (12)a reinforce the ideas of the author that you can learn more from traveling and shows the author's firm belief in this point. But AI uses less such thematic structures. It often writes the same sentence pattern such as listing example with "whether" clause: "Whether it's a job interview, a social event, or meeting someone for the first time, making a great first impression is crucial.". This is in line with the research of Ciordano Tudino and Yan Qin (2024): in terms of subordination, human-authored texts exhibit a significantly higher prevalence, indicating greater cohesion, coherence and the use of more complex linguistic constructs compared to ChatGPT's output.

(12)a, There is a Chinese proverb said, "Traveling thousands of miles is better than reading thousands of books. " It means you can learn more from traveling to places as opposed to reading about them from books. It is very true that you can learn more by actually experiencing things by yourself. (Text6 of human-written essays)

5. Conclusion:

This study compared the human-written essays with ChatGPT-generated essays through the analysis of thematic choices and thematic progression to check the coherence of them. From the statistics of topical Themes, human and ChatGPT both have a huge quantity of them and in the marked Themes, although they are more frequent in AI-generated articles, many of them are repeated, there are more various forms in human-written essays. In terms of interpersonal Themes, ChatGPT seldom adopts such

modal adjunct or modal verbal operators as interpersonal Themes which indicates that AI is lack of “writer awareness” to express his point of view and loses authenticity. In textual Themes, the machine uses conjunctive adjunct of addition and adversative to extend the clause and compare different situation and hardly uses other types of textual Themes but human employs more types of textual Themes to extend, combine and conclude the clauses, making the discourse logical, coherent and cohesive.

The analysis of thematic progression and coherence reveals that constant progression is the most frequent in ChatGPT-generated essays which shows that AI creates static texts and always repeats the same Themes. Human writers, on the other hand, use linear progression and derived progression more frequently and adopt distinctive ways to link clauses, following reader’s reading habits. Human-written essays have a clear information structure and discourse structure which make clauses interconnect with each other semantically and logically. It is more dynamic. But ChatGPT just made a list of his ideas instead of combining different thematic progressions. It uses single and simple thematic progression and even in a wrong way, making the texts inorganized and incoherent. Human also makes use of some special thematic structure to develop the discourse.

As the AI tool is continuing to update and develop, numerous people rely on its quick writing function. Developments in automated writing call for greater awareness on the part of L2 and writing instructors of what AI tools can provide in terms of writing assistance or authoring (Carvalho et al., 2022). The research of the study may have some instructions on the second or foreign language learning. Although ChatGPT can provide several ideas in a few seconds, the articles are not coherent, cohesive and complex enough and may make mistakes on organizing information. More importantly, it is lack of subjectivity and authenticity, creating tedious and inorganized essays. Therefore, when using ChatGPT, it is not suggested to fully copy its answer and students should identify its disadvantages and mistakes. And it is significant to make full use of thematic progression and to learn how to organize clauses and thus write a coherent essay.

Funding: This research received no external funding.

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

ORCID iD (if any)

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] Ariyaratne, S., Iyengar, P. K., Nischal, N., Babu, C. N., & Botchu, R. (2023). A comparison of ChatGPT-generated articles with human-written articles. *Skeletal Radiology*, 52, 1755-1758.
- [2] Dergaa, I., Chamari, K., Zmijewski, P., & Saad, B. H. (2023). From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, 40(2), 615-622.
- [3] Man, J. (2001). Thematic Progression, Semantic Cohesion and Coherence of English Writing-----Analysis of Model Essays for CET4 and 6. *Foreign Language Education*, 22(5), 45-50.
- [4] Sardinha, B. T. (2024). AI-generated vs human-authored texts: A multidimensional comparison. *Lingua*, 4(1), 1-10.
- [5] Thompson, G. (2014). *Introducing Functional Grammar*. Routledge.
- [6] Tudino, G., & Qin, Y. (2024). A corpus-driven comparative analysis of AI in academic discourse: Investigating ChatGPT-generated academic texts in social sciences. *Lingua*, 312(1), 1-20.
- [7] Xiong, M. Q., & Liu, X. M. (1999). Examining Several Coherence Theories in terms of Coherence Conditions. *Journal of Foreign Languages*, (3), 22-28.
- [8] Yang, L. X., & Ping, J. P. (2024). A Study of Generative AI in Academic Writing from the Perspective of Metadiscourse. *Contemporary Rhetoric*, (6), 11-23.
- [9] Yang, S., Chen, S., Zhu, H., Lin, J., & Wang, X. (2024). A comparative study of thematic choices and thematic progression patterns in human-written and AI-generated texts. *System*, 126(1), 1-15.
- [10] Zhang, D. L., & Liu, H. M. (1994). Thematic Structure and Textual Coherence. *Foreign Language Research*, (3), 27-33.
- [11] Zhao, P. (1998). Thematic Connection and Information Processing and Coherence of English Writing. *Foreign Language Research*, (1), 21-25.
- [12] Zhu, Y. S. (1995). Thematic Progression Pattern and Discourse Analysis. *Foreign Language Teaching and Research*, (3), 6-12.