



ARTICLE



<https://doi.org/10.1057/s41599-024-03726-7>

OPEN

What are the differences? A comparative study of generative artificial intelligence translation and human translation of scientific texts

Linling Fu¹✉ & Lei Liu¹

Generative artificial intelligence translation (GenAIT) brings convenience yet also imposes severe challenges on the dissemination of knowledge. The respective (dis)advantages of GenAIT and human translation (HT), and the ways to promote their effective interaction have not been sufficiently explored yet. This study investigates the linguistic features of GenAIT and HT of scientific texts rendered from English to Chinese from lexical and syntactic levels. The GenAIT is generated by ChatGPT 3.5, a representative GenAI platform, while HTs are done by 19 Master-of-Translation-and-Interpreting students in China. Data shows that GenAIT and HTs present distinguished linguistic features in both levels. At the lexical level, HT exhibits lengthier texts with a lower average word diversity; GenAIT presents higher accuracy in translating terminology. At the syntactic level, the average sentence count in HT is greater, whereas its average sentence length measured in tokens is shorter. Moreover, human translators tend to transform sentences from passive voice into active voice more frequently than ChatGPT 3.5 does. Furthermore, human translators exhibit superior skills in deconstructing lengthy and complex sentences into shorter, more comprehensible clauses. This study reveals that ChatGPT 3.5 and human translators exhibit complementary capabilities in the translation of scientific texts, thereby suggesting an optimisation of the strengths inherent to both parties. Implications are provided for future translator training, language service providers and the ongoing development of GenAIT and HT.

¹School of Foreign Studies, South China Normal University, Guangzhou, China. ✉email: llfu@scnu.edu.cn

Introduction

Generative artificial intelligence (GenAI) elevates machine translation (MT) into a new stage. Under the generative artificial intelligence translation (GenAIT) model, source texts can be translated and further revised by inputting prompts to an AI system like ChatGPT. Before the extensive application of GenAIT systems, Neural Machine Translation (NMT) is widely acknowledged for its notable performance in delivering translations with high accuracy and fluency (Klubička et al., 2017; Moorkens, 2018), especially in managing inversion in lexical level (Toral and Sánchez-Cartagena, 2017), improving sentence readability (Qin, 2018) and enhancing textual coherence (Zhang et al., 2020). In addition to its advantages, shortcomings of NMT have also been identified (Jiang and Niu, 2022). Nevertheless, to date, the linguistic features of GenAIT, and its (dis)advantages compared to human translation (HT) are still underexplored.

To fill this gap, a comparative study of GenAIT and HT has been conducted. In this study, ChatGPT 3.5 is employed to generate GenAIT and HTs are done by 19 MTI (Master of Translation and Interpreting) students in a university located in South China. The assumption underpinning this study is that GenAIT and HT present distinguished linguistic features and there is a possibility to integrate their respective strengths for higher quality translation. This empirical study focuses on the translation of scientific texts provided its significant role in promoting technological exchanges among countries, as highlighted by Cao (2000). Through its findings, the present study endeavours to enrich the understanding of GenAIT and facilitate effective collaboration between generative artificial intelligence and human translators.

Literature review

Linguistic features between MT and HT. GenAIT, the main concern of the present study, signifies an advanced phase in MT. Research in the comparison of MT-HT linguistic features can be mainly categorised into two types: one in the paradigm of translation studies and the other in the paradigm of linguistic analysis.

Research in the paradigm of translation studies has mainly focused on translation universals (Lapshinova-Koltunski, 2015; Krüger, 2020), translation quality (Ahrenberg, 2017; Muftah, 2022) and translationese (Loock, 2020). Drawing on translation universals (Baker, 1993), Lapshinova-Koltunski (2015) investigates the differences among HT, rule-based MT and statistical MT, concentrating on simplification, explicitation and convergence. Furthermore, Krüger (2020) delves into the explicitation in Neural Machine Translation (NMT), identifying that such translation shifts primarily occur through lexical insertion, lexical specification and relational specification. Regarding translation quality, BLEU and TER metrics are utilised in Ahrenberg (2017). Specifically, in terms of adequacy, Muftah (2022) finds no statistical difference between HT and MT from Google Translate and Babylon Translate. Regarding translationese, MT systems are effective in helping students to identify translation-specific language patterns (Loock, 2020).

Compared to the paradigm of translation studies, research in the paradigm of linguistic analysis has encompassed a broader range of topics, including lexical analysis (Klebanov and Flor, 2013; Vanmassenhove et al., 2019; Frankenberg-Garcia, 2022), morphologically complex words (Passban et al., 2018), cohesion (Wong and Kit, 2012; Niu et al., 2020), syntactic features (Sianipar and Sajarwa, 2021) and figurative language and style (Alowed, Al-Ahdal, 2023). Many scholars are interested in lexical analysis. For instance, the lexical tightness (a measure of association within a text) of HT has been found greater than that

of MT (Klebanov and Flor, 2013). However, when comparing MT with HT in English-French and English-Spanish language pairs, current MT systems fail to achieve the same level of lexical diversity as human translators do (Vanmassenhove et al., 2019). Moreover, through a corpus-driven keyword analysis, MT has problems like lexical inconsistency (Frankenberg-Garcia, 2022). With respect to morphological features, Passban et al. (2018) proposes a double-channel encoder and double-attentive decoder structure, enhancing the performance of NMT in translating morphologically complex words. Regarding cohesion, human translators tend to use more cohesive devices (Wong and Kit, 2012) and MT exhibits weaker coherence compared with HT (Niu et al., 2020). In terms of syntactic features, the use of passive voice, tense, sentence structure and voice conversion differs between MT and HT in Indonesian-English language pairs (Sianipar and Sajarwa, 2021). Concerning figurative language and style, AI-powered translation system is recognised as a powerful tool though it fails to properly translate the figure of speech and preserve the style of the source text (Alowed, Al-Ahdal, 2023). This aligns with Mehawesh's (2023) findings, which suggest that human translators are more effective in translating emotive words.

While previous studies in the field of MT-HT linguistic comparison have been extensive, there are notable gaps. Firstly, despite previous research on GenAIT-HT linguistic comparison among literary (Zagood et al., 2021) and legal texts (Moneus and Sahari, 2024), insufficient attention has been devoted to investigating linguistic differences between GenAIT and HT of scientific texts thus far. Secondly, primary emphasis within existing studies is mainly placed on Indo-European languages like French, English and Portuguese, with relatively few concerns on Chinese. With the rise of large language models (LLMs) like ChatGPT potentially offering alternatives to neural machine translation (Lionbridge.com, 2023), and given the vast number of Chinese speakers worldwide, further exploration of the linguistic attributes of GenAIT and HT within the context of Chinese-English scientific translation is imperative.

Scientific text translation. Scientific texts are informative, expository or referential, objectively ascertained and universally valid (Millan-Varela, 2013). As scientific texts offer insight into the development of specific fields (Dunham et al., 2020), studying translated scientific texts is indeed a meaningful way to facilitate the spread of knowledge.

Characteristics and translation standard of scientific texts should be clarified firstly. The scientific text, at lexical level, usually contains an array of proper nouns, abbreviations, tables, figures and so on. In syntactic dimension, the language of these texts is often complex, typified by extended sentence structures, a preference for the passive voice and a focus on noun phrases (Fu and Tang, 2012). Besides, the main function of scientific text is informative (Olohan, 2013). To fulfill the informative role of scientific text translation, maintaining accuracy and objectivity is paramount. However, securing these attributes in translations can pose a significant challenge for translators. For instance, an accurate and justifiable interpretation is a principal challenge for students in rendering English to Arabic (Abdulmajeed and Mohammed, 2023). Additionally, scientific texts often contain technical vocabulary, ordinary words with non-vernacular meanings, ellipsis, lengthy noun phrase and complex sentence (Fang, 2006).

Studies on the MT and HT of scientific texts span a diverse array of topics. For instance, Tehseen et al. (2018) proposes an effective scientific machine system which can work for any

domain. Besides, from the cognitive perspective, grammar is the main cause of most problems in translating a piece of scientific text into English (Minchenkov, 2019). As for the translation quality, Fan et al. (2020) demonstrates that computer-aided methods can generate translated scientific texts with quality closely resembling that of professional translators. Moreover, six error types in MT are identified through a fine-grained manual evaluation based on materials extracted from biomedical journals, biomedical books and so on (Liu et al., 2021).

However, studies focusing on the linguistic characteristics of GenAIT and HT within the realm of scientific texts are still scant. A comprehensive analysis of the linguistic elements in GenAIT and HT of scientific content is imperative to identify their respective strengths and weaknesses. Moreover, such an examination can facilitate the realisation of an effective symbiosis between the two.

Method

Research questions. To address the aforementioned gaps and investigate practical strategies for sustaining a productive interplay between GenAIT and HT, we design an empirical study engaging 19 Master-of-Translation-and-Interpreting (MTI) students. All the participants had been admitted to the same MTI programme in 2022 and none of them had prior experience in the translation industry before this task, indicating they were all student translators devoid of professional translation experience. GenAIT is generated by ChatGPT 3.5, a representative GenAI platform. Different linguistic features in lexical and syntactic level are supposed to be explored. The exclusion of textual level parameters from this study lies in two main reasons. Firstly, the limit of time for student translators to finish their translation tasks leads to the relatively short length of the source texts (ST) designed, and thereby the short length of target texts (TT) correspondingly. Limited text length results in unremarkable textual features, and thus textual features are not investigated as a research dimension in the present study. In addition, previous studies on the comparison of GenAIT-HT linguistic features have mainly focused on lexical and syntactic levels of the translated texts. Adhering to these two levels, we hope that our findings can verify those from previous literature more accordingly. This study attempts to answer the following two questions: In the translated scientific texts,

- (1) What are the lexical and syntactic features presented in HT and GenAIT respectively?
- (2) Based on these features, how to sustain the effective interaction between GenAIT and HT in scientific text translation?

Research design. Firstly, scientific texts’ translations generated by ChatGPT 3.5 and done by student translators are respectively collected. Then, a two-level analytical framework containing lexical and syntactic levels is built. Further, instruments of Wordless 2.3.0, CorpusWordParser 3.0, CLAWS-5 and AntConc 4.1.2 are used to collect data. Lastly, analysis of quantitative data of the linguistic features and qualitative analysis are implemented.

Materials and participants. This study involves 19 participants who are MTI students in a normal university of Southern China. All participants have signed informed consent voluntarily. Subsequently, they are tasked with two translation assignments, one from English to Chinese, and the other from Chinese to English, without obligatory sequence imposed on the completion of these tasks. The present study only focuses on the translation from English to Chinese as it is traditionally deemed that translators

maintain predominance in translating into their mother tongue (Newmark, 1981), and we tried to summarise features of HT from the language direction that translators are best at, which can avoid interference in the results due to their incompetence. Participants translate the text on computers without any online and offline auxiliary means, except for printed dictionaries. All participants have been told that their translations will be graded, thereby encouraging each participant to approach the task with diligence and finish it at their best.

The source texts chosen in this study are extracted from the teaching material of scientific and technological translation, published in authoritative publishing houses in China, ensuring the topic and difficulty of the ST are suitable for the participants. Before undertaking this translation task, none of the participants have used the aforementioned materials employed in this study, as it has not been incorporated into any of their coursework nor listed in the Recommended Books. Length of the text is identified in accordance with the workload-time ratio of the Level-2 of China Accreditation Test for Translators and Interpreters (CATTI), an authentic national translation proficiency test for which MTI students may target. CATTI Level-2 requires examinees to finish translating 500–600 English words into Chinese per hour. Given the time constraint of 1.5 h for the participants in finishing two translation tasks (one from English to Chinese and the other from Chinese to English), the English ST is designed to be 243 words in length. Thus, the 19 pieces of HT analysed in this study are English-to-Chinese translations completed by MTI students without the assistance of any electronic tools, and none of the HT texts have undergone post-editing.

The GenAIT collected in the present study is generated by ChatGPT 3.5. ChatGPT, as a large language model, requires prompts as guidance to function its translation ability (Jiao et al., 2023). The prompt utilised in our study is enlightened by Wang et al. (2023), which demonstrates the importance of context-aware prompts. In their study, the prompt “Translating this document from English to Chinese” is used to generate document translation. Consequently, in our study, the prompt “please translate the following passage into Chinese” has been employed. This prompt serves to inform ChatGPT that the ST is a passage and that the target language is supposed to be Chinese. The researchers input the prompt (operated on June 7th, 2023) into ChatGPT3.5 and paste the same English ST as that is translated by the 19 human translators, and thereupon the GenAIT is generated and collected. Additionally, this study includes a Reference Translated Text (RTT) for comparative analysis. This RTT is derived from the same source material as the original text (ST), ensuring consistency in content for a direct comparison.

Thus, the final corpus including ST, RTT, GenAIT and HT is collected. The general information is provided in Table 1 (The length is calculated by words in Microsoft Word 2019).

Framework of data analysis. At lexical level, the following parameters are observed: tokens, types, standardised type/token ratio (STTR), terminology and part of speech (POS). Token is

Table 1 General information of the final corpus.

	ST	RTT	GenAIT	HT
Language	English	Chinese	Chinese	Chinese
Length (words)	243	452	420	8795
Source	Teaching material	Teaching material	ChatGPT 3.5	19 MTI students

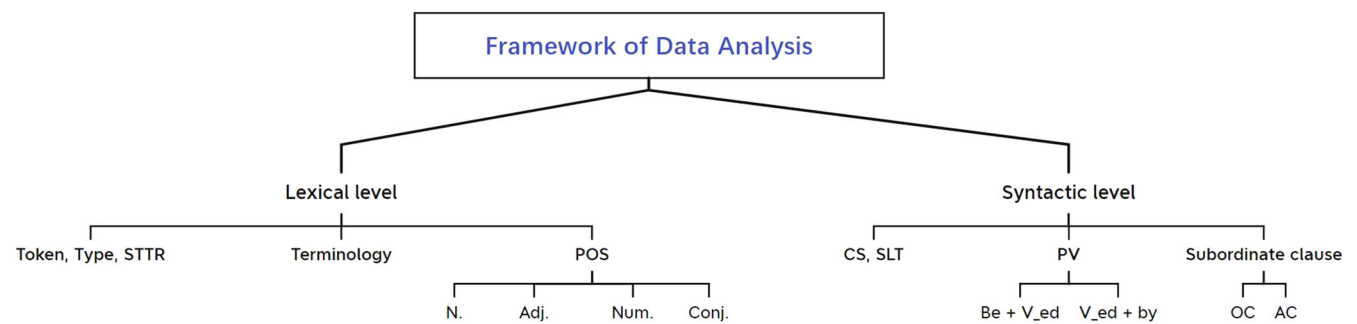


Fig. 1 Framework of data analysis.

defined as the total count of words within a text, which essentially represents the frequency of word occurrences (Kang and Yu, 2011; Bieswanger and Becker, 2017). Type, on the other hand, refers to the unique words that remain after eliminating repeated occurrences, essentially denoting non-redundant tokens (Kang and Yu, 2011; Liang et al., 2010). The Type-Token Ratio (TTR) is computed by dividing the total number of words in a text by the count of unique word types (Csomay and Crawford, 2024). The Standardised Type-Token Ratio (STTR) serves as an index to quantify lexical diversity within corpora (Calzada Pérez, 2017). According to Liang et al. (2010), STTR calculation involves a two-step process: initially determining the type/token ratio for each segment of 1000 words (adapted to 100 words in this study due to the brevity of the texts), followed by averaging these ratios to ascertain the STTR. In the present study, the STTR is automatically calculated through Wordless (Ye, 2023). Terminology is selected as a parameter because “scientific texts usually contain a variety of proper nouns” (Fu and Tang, 2012, p.2). Additionally, part of speech (POS) fulfills dual roles in both lexical and syntactic analyses. Given that our study prioritises word classification over sentence function when discussing the POS allocation, we categorise “part of speech” as a lexical feature. Based on the observation of the ST, nouns (N.) and adjectives (Adj.) accounts for a relatively large proportion, thus, they deserve to be noticed. Besides, the translation of numerals (Num.) is essential when a scientific text contains statistics. Given that using conjunctions (Conj.), as a type of connectors, is a major approach to enhance logical relations among clauses in scientific English material (Fu and Tang, 2012), conjunctions are included in the present study. Therefore, N., Adj., Num. and Conj. are chosen as sub-parameters of POS.

At syntactic level, the count of sentences (CS) and average length of a sentence (here as sentence length in tokens, SLT) are typical parameters to analyse syntactic features (Calzada Pérez, 2017). Besides, passive voice (PV) and subordinate clause, frequently used in scientific texts, are also included in the metrics. PV allows readers “to skim quickly and see the main points by looking at the subject of each sentence” (Biber and Conrad, 2009, p.123). Two types of PV structures, “be + V_{ed}” and “V_{ed} + by” (without “be”), are chosen in the analytical framework, for the former is an explicit structure and the latter can be regarded an implicit one. They are distinguished by whether the be-verb, as a marker of PV, appears in the structure. In this study, the two examples highlighted are “is called” and “led by”. Subordinate clause is included in the framework as “English scientific texts often involve various relations, premises and designated occasions, and clauses are abundantly existed” (Fu and Tang, 2012, p.86). Those adjective phrases used as attributes could be recognised as an abbreviation of attributive clause so as to strengthen the compact of the sentence structure (Randolph et al., 1985). Object clause (OC) and attributive clause are particularly selected for two reasons. First, it has been identified

Table 2 Parameters and methods/tools applied.	
Parameters	Methods/tools
Tokens, types, STTR, CS, SLT	Wordless 2.3.0
POS tagging (ST)	CLAWS-5
POS tagging (AIT, RTT and HT)	CorpusWordParser 3.0
POS extraction	AntConc 4.1.2
Terminology	Manually
PV	Manually
Subordinate Clause	Manually

that the clause type with the simple structure, “a nominal group + a verbal group + the other nominal group”, is frequently found in scientific text (Choi, 2013). This pattern is typical for OC. Second, the attributive clause accounts for a large proportion of long sentences in scientific English texts (Huang and Duan, 2021). Thus, an integrated two-level analytical framework is framed (Fig. 1.)

Tools and instruments. Four main tools are utilised for the extraction and analysis of the parameters listed in the analytical framework (Table 2). *Wordless 2.3.0*, developed by Lei Ye, is used to collect token, type, STTR, CS and SLT of the ST, RTT, GenAIT and HT. *Wordless* is embedded the function of detecting and coding various languages so that users neither need to manually set up the language coding of each file nor worry about the incompatibility due to unclear coding systems. *CLAWS-5*, a free online tool developed by Lancaster University, is used to tag POS of the ST, and *CorpusWordParser 3.0* is used for word segmentation and tagging of Chinese text. *AntConc 4.1.2* is utilised for the retrieval of POS. Thus, token, type, STTR, CS, SLT and POS are automatically collected while garbled codes or missing tags have been all corrected manually.

Researchers manually compile other parameters, including terminology, PV and subordinate clauses. The selection of terminologies from the ST is predicated upon their occurrences and their intrinsic significance to the context of the text. Subsequently, the translations corresponding to the selected terminologies from the RTT, GenAIT and HT are meticulously extracted. In all, ten cases of terminologies, two cases of PV structures, one object clause and one attributive clause are collected for analysis.

Results
Lexical features

Tokens, types and STTR. As there are 19 human translators in total, the mean value of data in human translations, HT (mean) is adopted. Additionally, we present tokens, types, and STTR of all HT in order to provide a whole picture of HT. HT (mean) out-numbers GenAIT in tokens and types while the STTRs of GenAIT and HT (mean) are nearly the same (0.659 and 0.650).

Specifically, all HTs, except HT05 (205) and HT15 (206), have more tokens than that of GenAIT. Besides, all HTs, except HT05 (133), have more types than that of GenAIT. However, around 68.42% of HTs have lower STTR than that of GenAIT (Table 3 and Figs. 2 and 3).

Table 3 reveals that tokens of the RTT are fewer than that of the ST, aligning with Baker’s (1993) postulated translation universal of simplification. The STTR is quantified through the type/token ratio, which endeavours to minimise the impact of functional words to the greatest extent. As such, STTR is employed as an indicator of lexical richness (Shi and Lei, 2021). Figure 3 shows that HT16 exhibits the minimum STTR value (0.583), signifying its potential inferiority in lexical richness when juxtaposed with that of GenAIT (0.66).

Translation of terminologies. GenAIT renders five out of ten terminologies (*hard discs, magnetic fields, current, reading and writing data, femtosecond*) into the same target expressions as RTT. In HTs, a superior congruency is observed: six out of ten terms (*retrieval, hard discs, IT, magnetic fields, current, reading and writing data*) are translated consistently with RTT, as delineated in Table 4.

Table 3 Tokens, types, and STTR descriptions (ST, RTT, GenAIT and HTs).				
Parameter	ST	RTT	GenAIT	HT (mean)
Tokens	246	229	208	236.68
Types	152	152	137	153.32
STTR	0.62	0.66	0.66	0.65

In general, HT exhibits advantages in translating abbreviated terms. There are 10 MTI students translating the term “IT” into “信息技术” (information technology) while GenAIT fails to do so. “IT” could be “互联网技术” (Internet technology), or “信息技术” (Information technology). Although there are overlaps between these two fields, “IT” in the ST means “information technology”, which justifies human translators outperforms GenAIT in dealing with abbreviated terms in this case.

However, GenAIT also performs its strength in the translation of “current”. The term “current” is translated into “现在” (right now) and “屏幕” (screen) in HT01 and HT03 respectively instead of the correct correspondence “电流” (current, related to electricity). Despite the term “current” encompassing all aforementioned meanings, the two MTI student translators fail to pick the correct one in the specific context yet ChatGPT 3.5 manages to do so.

POS. Table 5 illustrates that there are more nouns, adjectives, numerals, and conjunctions in HT (mean) than those of GenAIT. Further, except numerals, both GenAIT and HT present fewer counts of other three POS types than ST.

It also reveals that N. takes the largest proportion of all four POS types in ST (68.50%), RTT (76.03%), GenAIT (75.73%) and HT (mean) (73.95%). Furthermore, HT (mean) has more N. than GenAIT. For the quantities of Adj., Num. and Conj., the differences between GenAIT and HT (mean) are smaller. What’s more, the counts of Adj. in HT (mean) (10.68) and GenAIT (8) are within the scope of RTT (7) and ST (19). Num. sees the same feature. But for Conj., both GenAIT and HT (mean) presents fewer counts than ST and RTT.

Figure 4 shows that HT 16 and HT 02 contains 101 and 100 nouns respectively, ranked the first and second. In contrast, HT

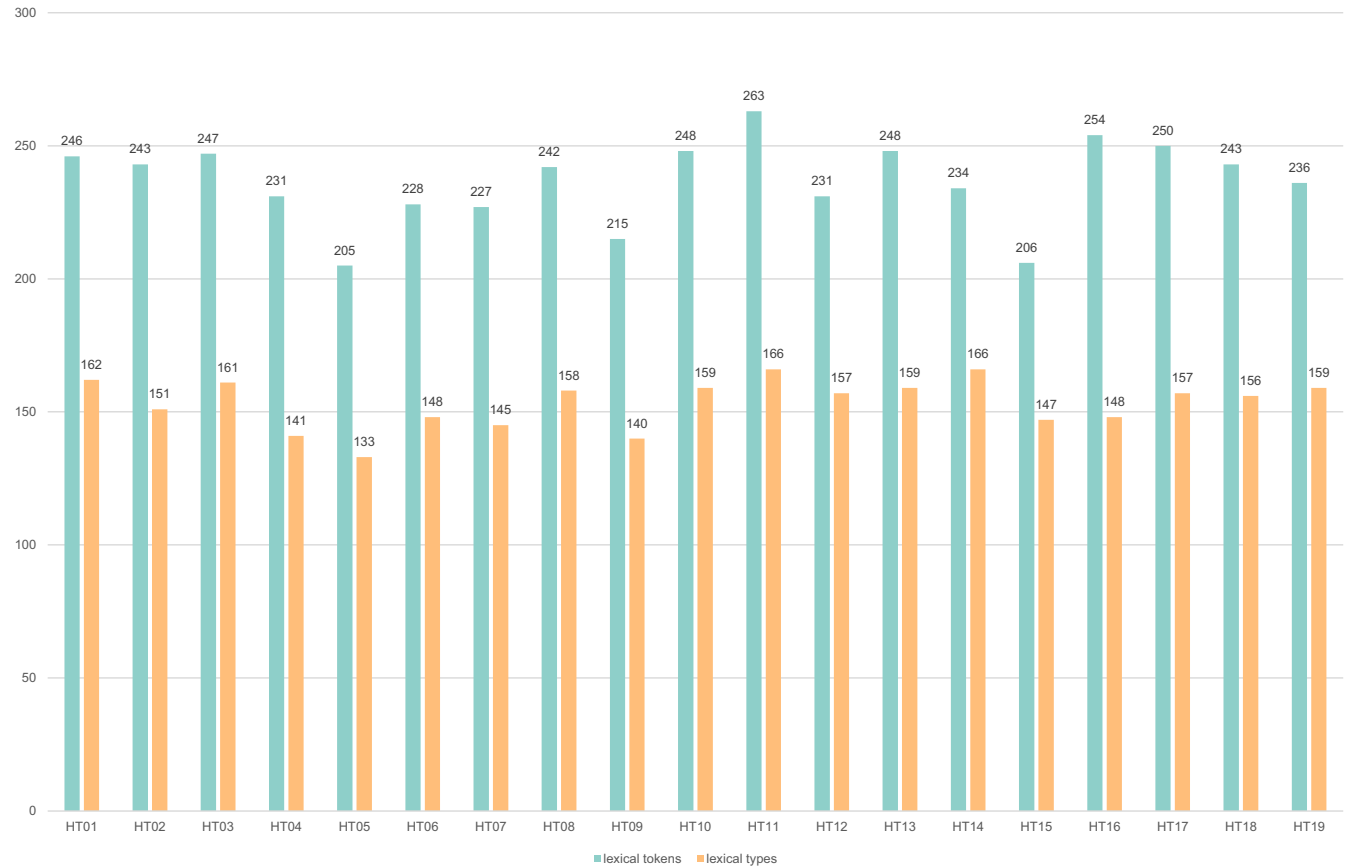


Fig. 2 Tokens and types of the 19 pieces of HT.

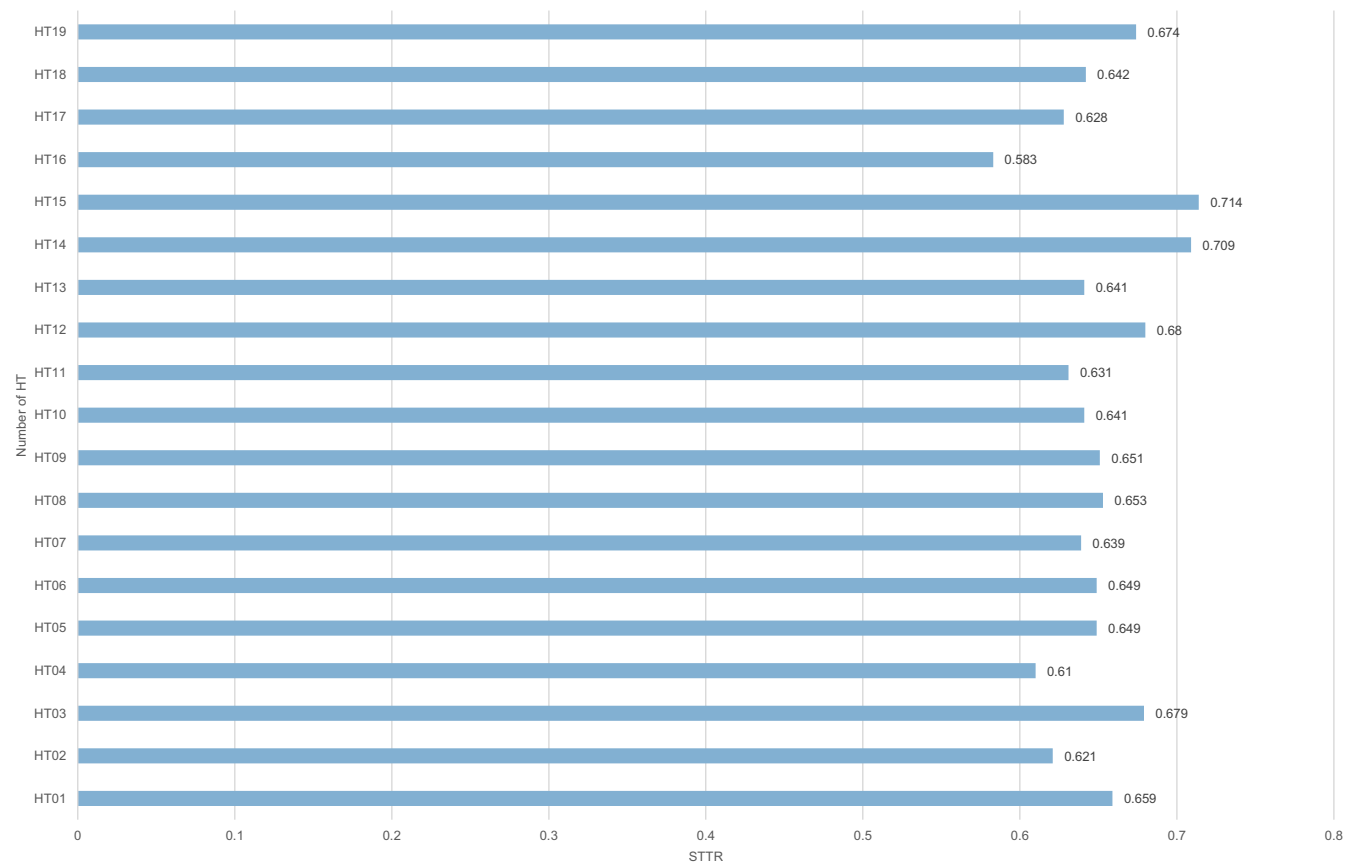


Fig. 3 STTR of the 19 pieces of HT.

Table 4 Ten selected terms and their translations in RTT, GenAIT and HT.						
No.	Term	RTT	GenAIT		HT	
			translations	Y/N	Counts/Y	Counts/N
1	Retrieval	检索	读取	N	10	9
2	Hard discs	硬盘	硬盘	Y	17	2
3	IT	信息技术	IT	N	10	9
4	Miniaturized storage	微型化处理器	微型存储	N	0	19
5	Magnetic fields	磁场	磁场	Y	17	2
6	Current	电流	电流	Y	15	4
7	Readout	读取器	读写	N	1	18
8	Reading and writing data	读写数据	读写数据	Y	11	8
9	Magnetic sensors	磁传感器	磁性传感器	N	3	17
10	Femtosecond	飞秒	飞秒	Y	2	17

*Note: "Y" means the translation is the same as RTT's. "N" means the translation is not exactly the same as RTT's.

Table 5 POS allocation in different types of translations.						
POS	ST (%)	RTT (%)	GenAIT (%)	HT		
				Mean	Sd.	Sd. ²
N.	87 (68.50)	92 (76.03)	78 (75.73)	85.32 (73.95%)	10.11	102.12
Adj.	19 (14.96)	7 (5.79)	8 (7.77)	10.68 (9.26%)	3.25	10.56
Num.	4 (3.15)	9 (7.44)	7 (6.79)	7.79 (6.75%)	1.69	2.84
Conj.	17 (13.39)	13 (10.74)	10 (9.71)	11.58 (10.04)	2.09	4.37
Total	127 (100)	121 (100)	103 (100)	115.37 (100%)	/	/

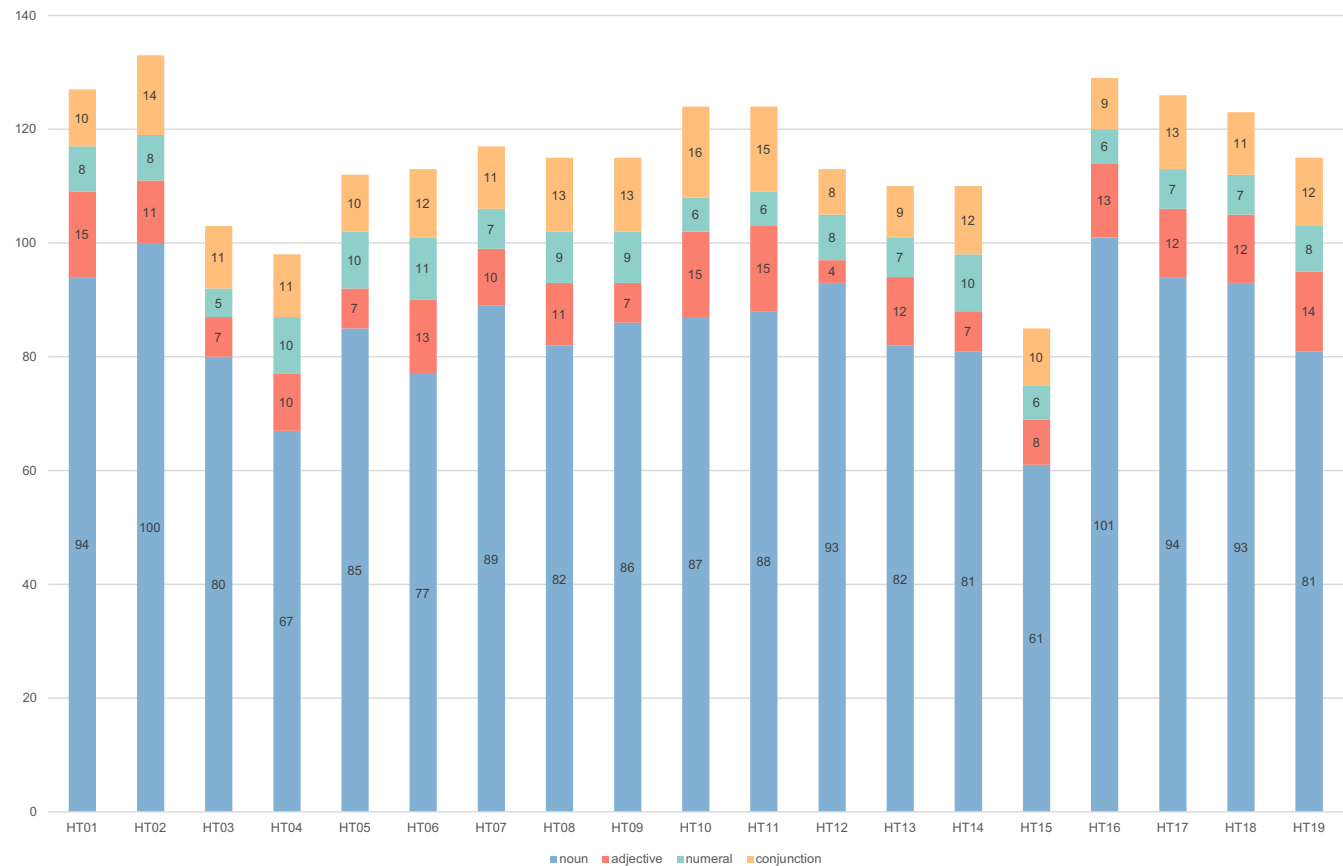


Fig. 4 POS of all HTs.

Table 6 Descriptive statistics of CS and SLT.						
Parameters	ST	RTT	GenAIT	HT		
				Mean	Sd.	Sd. ²
CS	9	10	9	11.68	1.97	3.89
SLT	27.33	22.60	23.11	20.71	3.42	11.72

15 has the fewest nouns (61). As for Num., the difference in all 19 pieces of HT is not statistically significant, ranges from five to eleven. 84.21% of HTs have 10 or more conjunctions. Additionally, 13 out of all the 19 pieces of HT contains over 10 conjunctions, exceeding the quantity of Conj. in GenAIT.

Syntactic features

Count of sentences and sentence length in tokens. The SLTs of RTT, GenAIT and HT (mean) are shorter than that of ST while the CSs of the ST, RTT, GenAIT and HT (mean) are diversified. Mean value of the CS in HT is higher than that of GenAIT, demonstrating that there are more sentences on average in HT (Table 6).

Nevertheless, there are still some exceptions. For example, HT03 has the fewest CS (8) while the largest SLT (29.625), which may indicate that the student cannot segment long sentences. For instance, the translated sentence “但通过自旋电子读取和写入数据却受到了发展相对滞后的电磁传感器的阻碍”(HT03, corresponding ST: But reading and writing data through spintronics has been hampered by the relative slowness of magnetic sensors.) is advised to be divided into several short sentences in Chinese.

Passive Voice (PV). Given that the proportion of PV in scientific writing might stabilise at about 30% (Ping Alvin, 2014), it deserves further explorations. This study focuses on two basic structures of PV: “be + V_ed”, an explicit one, and “V_ed + by” (without “be”), an implicit one.

As a fundamental PV structure, “be + V_ed” has been translated into different forms. RTT and HTs (except HT17) shift the PV into active voice (AV) in Chinese while GenAIT reserves the PV structure. “被” (the *bei* passive) is a widely used PV structure in Chinese. Nevertheless, rather than employing the *bei* passive construction, the OSV (Object-Subject-Verb) sequence is identified as the most similar counterpart to the English PV structure (Li, 2005), mainly because the *bei* passive is typically associated with adverse situations (Tsai, 1972; Li and Thompson, 1981). The example in Table 7 reflects that GenAIT should enhance its ability to recognise and properly translate English PV into Chinese. The student who does HT17 in Table 7 is also advised to pay more attention to the proper use of voice in Chinese.

Different from the explicit form of English PV “be + V_ed”, the structure of “V_ed + by” could be considered as an implicit one due to the lack of “be” as the sign. This example, if presented in its full form, could be “a team (which is) led by Jean-Yves Bigot”. In the example of Table 8, only HT07 shifts the PV into Chinese AV in the form of “you (have) structure” while GenAIT, like other HTs, seems to ignore the passive logical relation.

Subordinate clause. Due to the frequent use and large proportion of OC and attributive clause in scientific text (Randolph et al., 1985; Choi, 2013), this study selects one case for either type for analysis.

Table 7 Extracts of “be + V_ed” translated texts.	
Source	Text
ST	Our method is called the photonics of spin.
RTT	我们称之为自旋光子学。(Literal meaning: We call it the photonics of spin.)
GenAIT	我们的方法被称为自旋光子学。(Literal meaning: Our method is called the self -spin photonology.)
HT04	我们的方法称之为旋转的光子学。(Literal meaning: Our method calls it the photonics of spin.)
HT17	我们的方法被称为旋转的光子。(Literal meaning: Our method is called the spinning photonics.)

Table 8 Extracts of “V_ed+by” translated texts.	
Source	Text
ST	...a team led by Jean-Yves Bigot.
RTT由吉恩•伊维斯•比戈所带领的研究小组。(Literal meaning: The team which Jean-Yves Bigot leads)
GenAIT让-伊夫•比戈特领导的团队。(Literal meaning: The team which Jean-Yves Bigot leads.)
HT07由金姆•耶夫斯•比格 (Jean-Yves Bigot) 带领的团队。(Literal meaning: The team which has Jean-Yves Bigot as the leader.)

Table 9 Extracts of OC translated texts.	
Source	Text
ST	Fert and Grünberg discovered that tiny changes in magnetic fields can yield a large electric output.
RTT	费尔和格林贝格发现, 磁场的微弱变化可以产生大电流输出。(Literal meaning: Fer and Greenberg discovered, magnetic fields' weak changes can yield a large electric output.)
GenAIT	费尔特和格伦贝格发现微小的磁场变化可以产生巨大的电输出。(Literal meaning: Fert and Greenberg discovered tiny changes in magnetic fields can yield a large electric output.)
HT05	福特和哥伦伯格发现, 电磁场中微小的变化可以产出极大的电能。(Literal meaning: Fort and Greenberg discovered, magnetic fields' tiny changes can yield a large electric output.)

Table 10 Extracts of AC (restrictive) translated texts.	
Source	Text
ST	The research builds on achievements that earned the 2007 Nobel Physics Prize for Albert Fert of France and Peter Grünberg of Germany, who ushered in a revolution in miniaturized storage in the 1990s.
RTT	这项成果使他们获得了 2007年 的诺贝尔物理学奖。(Literal meaning: This achievement helped them earn the 2007 Nobel Prize in Physics)
GenAIT	两人因此获得了 2007年 的诺贝尔物理学奖。(Literal meaning: They two thus gained the 2007 Nobel Prize in Physics)
HT06	这一研究获得了 2017年 诺贝尔物理学奖, 得主是.....(Literal meaning: This research achieved the 2007 Nobel Prize in Physics , the owner is...)
The restrictive clauses are in bold type.	

Object clause (OC). In translating OC with “that” as the subordinating conjunction, RTT and most HTs (except HT01 and HT19) divide the complex sentence into two clauses with a comma while GenAIT does not (See Table 9). A plausible explanation might be that the source text lacks a comma, leading ChatGPT 3.5 to strictly adhere to the initial construction without alteration.

Attributive Clause (AC). The integrated sentence selected from ST for analysis is “The research builds on achievements that earned the 2007 Nobel Physics Prize for Albert Fert of France and Peter Grünberg of Germany, who ushered in a revolution in miniaturised storage in the 1990s.” The restrictive clause is chosen for the comparative analysis (See Table 10).

The restrictive clause forms a cause-and-effect relationship with the indefinite relative clause. The structure in GenAIT is similar to that of RTT, applying “因此” (thus) as a sign to show the connection between two clauses. On the contrary, some human translators rigidly follow the original structure (like HT06, whose way of connecting the two clauses is simply adding a comma), lacking logical connectors to manifest the relationship between two clauses.

Discussion

Lexical features

Lexical features in HT. Most HTs present larger number of tokens and types while lower STTR than those of GenAIT. Besides, compared with ST, there are 12 and 8 HTs using fewer tokens and types respectively. There are multiple factors to consider. Firstly, it could be the translation universal of simplification that results in over half (12/19) of HTs having fewer tokens. However, as for the six HTs (Chinese) that outnumber ST (English) in terms of tokens, it could be justified by that the information quantity of 1.5 Chinese characters is equivalent to one English word (Sun et al., 1985). Secondly, the linguistic differences between Chinese and English could also play a role. As those functional words (like “the”, “of”) are more frequently used in English, it affects the number of types and the STTR. Thirdly, the familiarity of individuals with the characteristics of scientific texts may also influence their translation results. Despite all 19 students being in the same class, it is inevitable that there will be disparity among them concerning their exposure to this specific genre.

In relation to the accuracy of terminology translation, human translations (HTs) demonstrate a lower proficiency. This may be

attributed to students' occasional carelessness or anxiety, given that they are working on a time-constrained translation task.

For part of speech (POS) analysis, the occurrence of nouns in HTs surpasses that of adjectives, numerals, and conjunctions, aligning with the expected lexical density and informational focus of scientific texts. Additionally, the standard deviation (Sd.) of numerals is the least, underscoring numerals' direct correlation with accuracy. The slight variation in numeral count across HTs may stem from differences in typing and segmentation practices. Furthermore, the relatively smaller Sd. of conjunctions suggests that the majority of students have effectively prioritised logical connectivity in their translations, adhering to findings that the incorporation of more conjunctions can enhance textual cohesion (Dalimunte, 2013).

Lexical features in GenAIT. Compared to HTs, GenAIT has fewer tokens and types, higher STTR, higher accuracy in terminology translation, lower word frequency of nouns, adjectives, numerals and conjunctions. These features are closely related to the working mechanism of ChatGPT. Trained on approximately 45 terabytes of text data at a multimillion-dollar cost (Barreto et al., 2023), GPT-3 sets the stage for its successor, ChatGPT 3.5. This subsequent model inherits GPT-3's extensive dataset but expands its capabilities with enhanced conversational interfaces that closely mimic human-like interactions (Rane, 2023), potentially augmenting its processing flexibility. Consequently, ChatGPT 3.5 might demonstrate an adeptness in adopting the stylistic features of scientific texts, which typically favour a higher frequency of content-driven words over functional ones. This attribute, combined with the nuanced training in linguistic variability, may explain the incremented STTR observed in ChatGPT 3.5's translations as compared to those produced by student translators. ChatGPT, to some extent, can be deemed as a unified multilingual machine translation model (Jiao et al., 2023). Although its architecture differs from the traditional encoder-decoder blueprint of NMT, GPT models, including ChatGPT 3.5, demonstrate promising translation capabilities (Hendy et al., 2023). This proficiency may be attributed to large language models' ability to abstract high-level cognitive capabilities through effective prompting in human-like conversations (Kojima et al., 2023), consequently enhancing the quality of text generated by ChatGPT. Furthermore, as a higher STTR signifies higher lexical diversity, it can be argued that ChatGPT 3.5 yields translations exhibiting broader vocabulary variance compared with student translators in this study. This finding contrasts with Vanmassenhove et al. (2019) which shows a decrease in lexical diversity with MT. Nevertheless, it's pertinent to note that the system utilised in our study is different from the MT assessed by Vanmassenhove et al. (2019).

As for the count of POS, except the frequency of nouns, counts of the other three POS in GenAIT are similar to those of ST. The reason for the higher frequency of nouns might be that ChatGPT 3.5 transforms many English verbs into Chinese nouns. For instance, the verb "output" is translated into "输出" (production).

Syntactic features

Syntactic features in HT. HTs exhibit a larger CS coupled with shorter average sentence lengths in comparison to GenAIT. Student translators' translation training might explain this result. When Chinese students embark on English-Chinese translation tasks, on translating long English sentences in particular, they tend to split them into several short sentences. As one of common translation strategies, it has been consolidated through abundant training, and students tend to apply this strategy to practice.

Regarding the translation of PV, most MTI students transform PV into AV. Specifically, all student translators employ the AV structure in Chinese when translating the implicit PV structure "V_{ed} + by". This might be associated with a prevalent use of the "有(you, have)-structure" in Chinese language contexts. However, whether to transform PV into AV is influenced by various reasons. For example, a study of English-German reveals that translators mainly use PV to translate AV but editors change most of PV back to AV (Bisiada, 2019).

The translation of OC and AC reflects student translators' advantages in segmenting long sentences. OC "most frequently appear after verbs like say, believe, think, decide, propose... and the verbs that name some sort of mental activity" (Podoliuc, 2010, p.51). These complements are instrumental in constructing complex sentences. When encountering such complex constructions, student translators opt to divide them into two distinct parts using the connector "that".

Syntactic features in GenAIT. The GenAIT has fewer sentences (9) than HTs' average (11.68) but larger SLT (23.11) than that of HTs (20.71). To some degree, the more tokens in a sentence, the lower readability the sentence is. Although readability should be assessed from various perspectives, including syntax, morphology, cohesion, discourse structure and subject matter (François and Fairon, 2012; Denning et al., 2016; Arfé et al., 2018), calculating average sentence length or syllables per word is still a prominent strategy (Azpiazu and Pera, 2019). As there are more tokens in sentences of GenAIT than those of HTs on average, readers might need more efforts when reading the GenAIT.

In relation to translations of PV in GenAIT, the explicit PV structure "be + V_{ed}" is remained while the passive relation marked by the implicit "V_{ed} + by" seems to be ignored. The "be + V_{ed}" structure unambiguously signals that the action is executed by someone, thereby prompting ChatGPT to preserve the passive construction rather than converting the PV to AV. However, with more implicit structures, ChatGPT may struggle to detect the passive construction, leading to potential oversight.

Regarding subordinate clause, GenAIT maintains object complements as intact, whereas most human translators elect to divide them. ChatGPT 3.5 employs a connector to seamlessly join the restrictive attributive clause with the indefinite relative clause, thereby enhancing the clarity of the source text for readers. Similarly, translators of HT01, HT03 and HT19 do not split the OC either. Although NMT systems generally exhibit limitations in generating coherent logical information (Wang and Feng, 2021), the use of connectors by GenAIT demonstrates its partial competency in constructing logical connections. Notably, the ST lacks conjunctions such as "so" or "therefore" to bridge the restrictive attributive clause and the indefinite relative clause. In contrast, ChatGPT 3.5 introduces "因此" (thus) to forge this link.

The ways to sustain the effective GenAIT-HT interaction. Based on above findings, harnessing the lexical and syntactic advantages of both the GenAIT and HT could be an effective interaction between them to improve translation quality. GenAIT offers human translators a wide range of vocabulary and a wealth of background knowledge across various fields. Human translators can effectively leverage words with different parts of speech, allowing them to combine the strengths of both sides and enhance the lexical accuracy of the translated text. For instance, when human translators encounter engineering terminologies that are unfamiliar to them, they could refer to the knowledge provided by GenAI platform to achieve more precise translations. Conversely, when human translators identify inappropriate word usage in GenAIT, such as words belonging to improper parts of speech, they could make adjustments to enhance the accuracy of the translated text.

At syntactic level, the effective interaction of GenAIT and HT can enhance the fluency of translated texts. In this study, student translators excel in deconstructing long and complex sentences. They can segment a lengthy passage into manageable sections for sequential input into GenAI. This strategy facilitates the improvement of GenAI's output by catering to its preference for processing shorter and simpler sentences. For example, in the case of translating a long poem written in archaic English, a task that remains a formidable challenge for current AI technologies, human translators can strategically divide it into segments based on its meaning and context. Concurrently, GenAI's capabilities, like the automatic insertion of logical connectors in this research, can further assist human translators in augmenting the fluency of their target texts.

Implications

Drawing from the findings of this research, we propose several implications for translator training, language service providers, and future development of GenAI and human translation. In terms of translator training, educators are encouraged to guide students in effectively utilising GenAI systems to improve their translation quality. For instance, given that GenAI systems like ChatGPT in this study display a higher STTR which indicates richer lexical diversity, it is advisable for trainers to foster students' proficiency in exploiting GenAI tools to diversify their lexical choices. Additionally, given ChatGPT's superior accuracy of translating numerals in this research, it is suggested that student translators could employ GenAI systems to verify numerical translations and thus enhance both the efficiency and quality of their work. Nonetheless, translators must be vigilant in scrutinising the input material processed through GenAI platforms, especially long sentences, as GenAI is found to be less effective in dealing with long sentences.

For language service providers (LSPs), it is highly recommended to deepen their understanding and optimise the efficient utilisation of GenAI technologies. For example, in the present study, ChatGPT 3.5 demonstrates advanced lexical diversity and effective clause connection, which could substantially enhance efficiency within LSP environments. Leveraging these advantages of GenAI has the potential to result in providing more sophisticated and nuanced language services. Besides, in September 2023, Lionbridge, a leading language service provider, published its Machine Translation Report, highlighting two pivotal insights. Firstly, companies that successfully integrate AI will likely excel in the evolving digital landscape. Secondly, large language models like ChatGPT could emerge as a potential successor to neural NMT as a new standard in the field. Consequently, it is recommended that LSPs invest in the development of GenAI and in equipping their workforce with AI-enhanced skills to further boost productivity.

The future landscape for GenAI and human translators is supposed to be a harmonious coexistence with a bright outlook. It's undeniable that GenAIT has shown notable advantages in certain aspects. Nonetheless, widespread consensus maintains that human translators are irreplaceable, as GenAI systems like ChatGPT have yet to adeptly capture uniquely human imaginative skills, such as creativity and emotional intelligence (Xu and Yang, 2022; Hu and Li, 2016). Moreover, this study also underscores GenAIT's limitations, including difficulties with processing lengthy sentences. As such, GenAI should not be viewed as a threat but rather as a valuable tool that enhances our capabilities. To utilise GenAI more effectively, there's a growing recognition of the need for "prompt engineering", a skill that demands specialised knowledge and practice (Oppenlaender et al., 2023).

Conclusion

Major findings. For the first research question, HT and GenAIT present distinguished lexical and syntactic features. In lexical level, HT (mean) has larger number of tokens and types but lower STTR than those of GenAIT; GenAIT has higher accuracy in term translation, fewer nouns, adjectives, numerals and conjunctions. In syntactic level, HT (mean) has more sentences and smaller SLT on average. Most HTs (except HT17) transforms the explicit PV structure into AV in Chinese while GenAIT fails to do so. The implicit PV structure is shifted into AV in all HTs while ChatGPT 3.5 seems to ignore the passive relation. Though ChatGPT 3.5 fails to split the complex sentence into simple sentences, it manifests the cause-and-effect relation of the two attributive clauses by adding a connector “因此” (thus) as some student translators do. For the second research question, combining advantages of both sides is proposed to be a method of sustaining the effective interaction between GenAIT and HT.

Limitations and future directions. Different from the previous study on perceptions and attitudes (Liu et al., 2022), this study provided objective evidence for the complementarities between GenAIT and HT. Thus, this research represents a modest progress. Besides, based on findings, this study points out specific aspects where GenAIT and HT could benefit each other. Implications for translator training, language service provision and the development of GenAIT and HT are also demonstrated. However, some limitations must be noted. Although parameters in the analytical framework are all justified, there are many other dimensions deserved to be investigated. For example, if the material is extracted from a storytelling book, probably verbs should be added into the POS in the framework as verbs might frequently appear in such texts. In addition, this study ignores textual level due to it is not prominent in the selected material. Future researchers are supposed to explore this dimension and involve a larger number of participants. Besides, other language pairs should be researched to see if these findings are replicable in different language contexts. As for other text genres, legal text, literature and many other genres are also deserved to be explored for a more extensive application.

Data availability

All relevant data of this study are publicly available at Open Science Framework (https://osf.io/gxam6/?view_only=da0cdabf22f46d195f224ced9121977).

Received: 8 December 2023; Accepted: 5 September 2024;

Published online: 17 September 2024

References

- Abdulmajeed AN, Mohammed LN (2023) Linguistic features of scientific texts in translation. *Adab Al-Rafidayn* 53(92):76–89. <https://doi.org/10.33899/radab.2023.177616>
- Ahrenberg L (2017) Comparing Machine Translation and Human Translation: A Case Study. In: *Proceedings of the Workshop Human-Informed Translation and Interpreting Technology*, Association for Computational Linguistics, Shoumen, Bulgaria, pp 21–28. https://doi.org/10.26615/978-954-452-042-7_003
- Alowedi NA, Al-Ahdal AAMH (2023) Artificial Intelligence based Arabic-to-English machine versus human translation of poetry: an analytical study of outcomes. *J Namib Stud* 33:1523–1538. <https://doi.org/10.59670/jns.v33i.800>
- Arfé B, Mason L, Fajardo I (2018) Simplifying informational text structure for struggling readers. *Read Writ* 31(9):2191–2210. <https://doi.org/10.1007/s11445-017-9785-6>

- Azpiazu IM, Pera MS (2019) Multiattentive recurrent neural network architecture for multilingual readability assessment. *Trans Assoc Comput Linguist* 7:421–436. https://doi.org/10.1162/tacl_a_00278
- Baker M (1993) Corpus Linguistics and Translation Studies — Implications and Applications. In: Baker M, Francis G, Tognini-Bonelli E (eds) *Text and technology*. John Benjamins Publishing Company, Philadelphia, p 233–252. <https://doi.org/10.1075/z.64.15bak>
- Barreto F, Moharkar L, Shirodkar M, Sarode V, Gonsalves S, Johns A (2023) Generative artificial intelligence: Opportunities and challenges of large language models. Balas VE, Semwal VB, Khandare A (eds) *Intelligent computing and networking*. IC-ICN 2023, Springer, Singapore, p 545–553. https://doi.org/10.1007/978-981-99-3177-4_41
- Biber D, Conrad S (2009) *Register, genre, and style*. Cambridge University Press
- Bieswanger M, Becker A (2017) *Introduction to English linguistics*. De Gruyter Mouton, Boston
- Bisiada M (2019) Translated language or edited language? A study of passive constructions in translation manuscripts and their published versions. *Across Lang Cult* 20(1):35–56. <https://doi.org/10.1556/084.2019.20.1.2>
- Calzada Pérez M (2017) Corpus-based methods for Comparative Translation and Interpreting Studies: mapping differences and similarities with traditional and innovative tools. *Transl Interpret Stud* 12(2):231–252. <https://doi.org/10.1075/tis.12.2.03cal>
- Cao GY (2000) 科技英语文献的翻译研究 [Research on the Translation of EST Literature]. *Chin Sci Technol Transl* 1:8–11
- Choi GH (2013) Translating grammatical metaphor in a scientific text: Implications of understanding ‘favourite clause type’ for translator training. *T&I Review* 3:7–38
- Csomay E, Crawford WJ (2024) *Doing corpus linguistics*. Routledge, New York
- Dalimunte MHA (2013) English Metafunction analysis in chemistry text: characterization of scientific text. *Int J Engl Lang Trans Stud* 1(2):52–68
- Denning J, Pera MS, Ng YK (2016) A readability level prediction tool for K-12 books. *J Assoc Inf Sci Technol* 67(3):550–565. <https://doi.org/10.1002/asi.23417>
- Dunham J, Melot J, Murdick D (2020) Identifying the development and application of artificial intelligence in scientific text. *arXiv*. <https://doi.org/10.48550/arXiv.2002.07143>
- Fan D, Chen R, Xue J, Zhao Y (2020) Quality classification and evaluation of human-machine composite translations of scientific text based on KPCA. In: 2020 IEEE 3rd International Conference on Computer and Communication Engineering Technology (CCET), IEEE, Beijing, pp 163–166. <https://doi.org/10.1109/CCET50901.2020.9213087>
- Fang Z (2006) The language demands of science reading in middle school. *Int J Sci Educ* 28(5):491–520. <https://doi.org/10.1080/09500690500339092>
- François T, Fairon C (2012) An “AI readability” formula for French as a foreign language. In: *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, Association for Computational Linguistics, USA, pp 466–477. <https://doi.org/10.5555/2390948.2391004>
- Frankenberg-Garcia A (2022) Can a corpus-driven lexical analysis of human and machine translation unveil discourse features that set them apart? *Target* 34(2):278–308. <https://doi.org/10.1075/target.20065.fra>
- Fu YL, Tang YQ (2012) 科技翻译 [Scientific Translation]. Foreign Language Teaching and Research Press, Beijing
- Hendy A, Abdelrehim M, Sharaf A, Raunak V, Gabr M, Matsushita H, Kim YJ, Afify M, Awadalla HH (2023) How good are gpt models at machine translation? a comprehensive evaluation. *arXiv*. <https://doi.org/10.48550/arXiv.2302.09210>
- Hu KB, Li Y (2016) 机器翻译特征及其与人工翻译关系的研究 [a study of characteristics in machine translation and its relationship with human translation]. *Chin Transl J* 37:10–14
- Huang HY, Duan HQ (2021) 科技英语定语从句的机译评价 [on machine translation evaluation in translating attributive clauses in scientific english text]. *Chin Sci Technol Transl* 34:21–24
- Jiang Y, Niu J (2022) A corpus-based search for machine translationese in terms of discourse coherence. *Across Lang Cult* 23(2):148–166. <https://doi.org/10.1556/084.2022.00182>
- Jiao W, Wang W, Huang JT, Wang X, Tu Z (2023) Is ChatGPT a good translator? A preliminary study. <https://arxiv.org/abs/2301.08745>
- Kang N, Yu Q (2011) Corpus-based stylistic analysis of tourism English. *J Lang Teach Res* 2(1):129–136. <https://doi.org/10.4304/jltr.2.1.129-136>
- Klebanov BB, Flor M (2013) Associative texture is lost in translation. In: *Proceedings of the Workshop on Discourse in Machine Translation*, Association for Computational Linguistics, Sofia, Bulgaria, pp 27–32. <https://aclanthology.org/W13-3304>
- Klubička F, Toral A, Sánchez-Cartagena VM (2017) Fine-grained human evaluation of neural versus phrase-based machine translation. *Prague Bull Math Linguist* 108(1):121–132. <https://doi.org/10.1515/pralin-2017-0014>
- Kojima T, Gu SS, Reid M, Matsuo Y, Iwasawa Y (2023) Large language models are zero-shot reasoners. In: *Advances in Neural Information Processing Systems* 35, NeurIPS, San Diego, USA, pp 22199–22213
- Krüger R (2020) Explicitation in neural machine translation. *Across Lang Cult* 21(2):195–216. <https://doi.org/10.1556/084.2020.00012>
- Lapshinova-Koltunski E (2015) Variation in translation: evidence from corpora. Claudio Fantinuoli C, Federico Zanettin F (eds) *New directions in corpus-based translation studies*, Language Science Press, Berlin, Germany, pp 79–99
- Li CN, Thompson SA (1981) *Mandarin Chinese: A functional reference grammar*. University of California Press, California
- Li ES (2005) Voice in Chinese: a systemic functional perspective. *Funct Lang* 12(2):181–203. <https://doi.org/10.1075/fol.12.2.03sum>
- Liang MC, Li WZ, Xu JJ (2010) *Using corpora: A practical coursebook*. Foreign Language Teaching and Research Press, Beijing
- Lionbridge.com (2023) The Lionbridge: 2023 machine translation report. <https://www.lionbridge.com/content/dam/lionbridge/pages/whitepapers/the-lionbridge-2023-machine-translation-report/lionbridge-2023-machine-translation-report-whitepaper.pdf>. Accessed 31 Jan 2024
- Liu K, Kwok HL, Liu J, Cheung AKF (2022) Sustainability and influence of machine translation: perceptions and attitudes of translation instructors and learners in Hong Kong. *Sustainability* 14(11):6399. <https://doi.org/10.3390/su14116399>
- Liu M, Zhang H, Wu G (2021) Fine grained human evaluation for English-to-Chinese machine translation: a case study on scientific text. *arXiv*. <https://doi.org/10.48550/arXiv.2110.14766>
- Look R (2020) No more rage against the machine: how the corpus-based identification of machine-translationese can lead to student empowerment. *JoS-Trans* 34:150–170
- Mehawesh MI (2023) Human translation vs. Machine translation in the Naguib Mahfouz’s novel palace walk: a case study. *MJHSS* 38(2). <https://doi.org/10.35682/mjhsc.v38i2.631>
- Millan-Varela C (2013) *The Routledge handbook of translation studies*. Routledge, London
- Minchenkov A (2019) Translating a scientific text into English: Cognitive perspective. *J Teach Engl Specif Acad Purp* 7(1):71–84. <https://doi.org/10.22190/jtesap1901071m>
- Moneus AM, Sahari Y (2024) Artificial intelligence and human translation: a contrastive study based on legal texts. *Heliyon*, 10(6). <https://doi.org/10.1016/j.heliyon.2024.e28106>
- Moorkens J (2018) What to expect from Neural Machine Translation: a practical in-class translation evaluation exercise. *Interpret Transl Train* 12(4):375–387. <https://doi.org/10.1080/1750399x.2018.1501639>
- Muftah M (2022) Machine vs human translation: a new reality or a threat to professional Arabic–English translators. *PSU Research Review*. <https://doi.org/10.1108/prr-02-2022-0024>
- Newmark P (1981) *Approaches to translation*. Pergamon Press, Oxford
- Niu J, Jiang Y, Zhou Y (2020) Approaching textual coherence of machine translation with complex network. *Int J Mod Phys C* 31(12):2050175. <https://doi.org/10.1142/s0129183120501752>
- Olohan M (2013) *Scientific and technical translation*. Routledge, London & New York
- Oppenlaender J, Linder R, Silvennoinen J (2023) Prompting AI art: An investigation into the creative skill of prompt engineering. *arXiv*. <https://doi.org/10.48550/arXiv.2303.13534>
- Passban P, Way A, Liu Q (2018) Tailoring neural architectures for translating from morphologically rich languages. In: 27th International Conference on Computational Linguistics, Santa Fe, New Mexico, USA, pp 20–26. <https://aclanthology.org/C18-1265>
- Ping Alvin L (2014) The passive voice in scientific writing. The current norm in science journals. *J Sci Commun* 13(1):1–16. <https://doi.org/10.22323/2.13010203>
- Podoliuc T (2010) Difficulties of the word “That” in the English complex sentences. *Intertext* (3–4):48–55
- Qin Y (2018) 基于神经网络的机器翻译质量评析及对翻译教学的影响 [An analytical study of neural network machine translation and its impacts on translator training]. *Foreign Lang Ed* 2:51–56
- Randolph Q, Greenbaum S, Leech G, Svartvik J (1985) *A comprehensive grammar of the english language*. Longman, London
- Rane N (2023) Role and challenges of ChatGPT and similar generative artificial intelligence in arts and humanities. *SSRN Electron J*. <https://doi.org/10.2139/ssrn.4603208>
- Shi Y, Lei L (2021) Lexical use and social class: a study on lexical richness, word length, and word class in spoken English. *Lingua* 262:103155. <https://doi.org/10.1016/j.lingua.2021.103155>
- Sianipar IF, Sajarwa S (2021) The translation of Indonesian passive voice in research articles’ abstracts into English: Human vs machine translation. *Ling Didakt J Bhs Dan Pembelajaran Bhs* 15(2):136–148. <https://doi.org/10.24036/ld.v15i2.112967>
- Sun F, Morita M, Stark LW (1985) Comparative patterns of reading eye movement in Chinese and English. *Percept Psychol* 37(6):502–506. <https://doi.org/10.3758/bf03204913>

- Tehseen I, Tahir GR, Shakeel K, Ali M (2018) Corpus Based Machine Translation for Scientific Text. In: Iliadis L, Maglogiannis I, Plagianakos V (eds) *Artificial Intelligence Applications and Innovations. IFIP Advances in Information and Communication Technology*, vol 519. Springer, Cham, Switzerland, pp 196–206
- Toral A, Sánchez-Cartagena VM (2017) A multifaceted evaluation of neural versus phrase-based machine translation for 9 language directions. arXiv. <https://doi.org/10.48550/arXiv.1701.02901>
- Tsai SG (1972) A study of translation. You Lian Publishing Company, Hong Kong
- Vanmassenhove E, Shterionov D, Way A (2019) Lost in translation: Loss and decay of linguistic richness in machine translation. arXiv. <https://doi.org/10.48550/arXiv.1906.12068>
- Wang L, Lyu C, Ji T, Zhang Z, Yu D, Shi S, Tu Z (2023) Document-level machine translation with large language models. arXiv. <https://doi.org/10.48550/arXiv.2304.02210>
- Wang W, Feng A (2021) Self-information loss compensation learning for machine-generated text detection. *Math Probl Eng* 2021:1–7. <https://doi.org/10.1155/2021/6669468>
- Wong BT, Kit C (2012) Extending machine translation evaluation metrics with lexical cohesion to document level. In: *Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning*. Association for Computational Linguistics, Jeju Island, Korea, pp 1060–1068. <https://aclanthology.org/D12-1097>
- Xu X, Yang Z (2022) The use of deep learning and AIoT technology in loan word translation. *Int J Grid Util Comput* 13(1):49–56. <https://doi.org/10.1504/ijguc.2022.10045601>
- Ye L (2023) Wordless (Version 2.3.0) [Computer software]. Github. <https://github.com/BLKSerene/Wordless>. Accessed 31 Jan 2024
- Zagood MJ, Al-Nuaimi A, Al-Blooshi A (2021) Man vs machine: a comparison of linguistic, cultural, and stylistic levels in literary translation. *Int J Linguist, Lit Transl* 4(2):70–77. <https://doi.org/10.32996/ijlt.2021.4.2.10>
- Zhang J, Utiyama M, Sumita E, Neubig G, Nakamura S (2020) Improving neural machine translation through phrase-based soft forced decoding. *Mach Transl* 34(1):21–39. <https://doi.org/10.1007/s10590-020-09244-y>

Acknowledgements

This research was supported by the National Social Science Fund of China (23FYYB041), the Guangdong Philosophy and Social Science Fund (GD22WZX02-04), the Center for Translation Studies of Guangdong University of Foreign Studies (CTS202001) and the Scientific Research Innovation Project of Graduate School of South China Normal University (2024KYLX029).

Author contributions

Conception and design of the work: LF. Supervision: LF. Original draft: LF and LL. Revising and editing: LF and LL. Data analysis and interpretation: LF and LL.

Competing interests

The authors declare no competing interests.

Ethical approval

Ethical approval was obtained from the Human Research Ethics Committee of the School of Foreign Studies at South China Normal University. It was confirmed that the research complied with ethical standards and was performed in accordance with relevant guidelines/regulations.

Informed consent

Informed consent was obtained from all participants in the study. We have explained to participants the detailed procedure, how data would be anonymised, and how their privacy would be protected. We assured participants that they could withdraw from the task at any time without facing any consequences. The participants were all adults, and their participation was entirely voluntary.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-024-03726-7>.

Correspondence and requests for materials should be addressed to Linling Fu.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024