Lexical Cohesion in English-Chinese Business Translation: Human Translators Versus ChatGPT

Na Tang[1], & Mohamed Abdou Moindjie[2]

[1] Phd Candidate, School of Languages, Literacies and Translation, Universiti Sains Malaysia, Malaysia

[2] Senior Lecturer, School of Languages, Literacies and Translation, Universiti Sains Malaysia, Malaysia

Correspondence: Na Tang, Phd Candidate, School of Languages, Literacies and Translation, Universiti Sains Malaysia, Malaysia.

Received: July 23, 2024 Accepted: September 26, 2024 Online Published: December 12, 2024

doi:10.5430/wjel.v15n2p286 URL: https://doi.org/10.5430/wjel.v15n2p286

Abstract

Lexical cohesion involves the continuity of text on the level of lexis achieved through word choices; it embodies repetition, synonymy, and collocation (Halliday, 1985). This study attempts to elucidate the lexical cohesion and examine translation shifts in the English-Chinese translation of business texts. The English business texts are compared with two Chinese translations: one by human translators and the other by ChatGPT. The research is based on Halliday's (1985) cohesion and Toury's (2012) descriptive translation studies. The data of lexical cohesion are identified and collected manually in a parallel corpus. The analysis deals with the description of lexical cohesion and translation shifts. The research reveals that semantic meanings of the items of lexical cohesion are largely maintained in the English-Chinese business translation. Additionally, despite using translation methods like literal translation, addition, omission, and conversion in translating lexical cohesion, human translators make more translation shifts of lexical cohesion than ChatGPT.

Keywords: lexical cohesion, English-Chinese translation, business translation, human translation, ChatGPT's translation

1. Introduction

The demand for business translation is large, yet it remains challenging (Chiper, 2002). Business translation surpasses literary translation in both quantity and economic value (Gotti & Šarčević, 2006). The need for business translation between English and Chinese is growing along with the increasing trade between China and the West in the context of globalization. Since the implementation of the Reform and Opening-Up Policy in 1978, the language pair of English-Chinese has supplanted Russian-Chinese as the dominant demand generator in the Chinese translation market (Chan, 2017). However, business translation between English and Chinese is challenging, not only because they originate from distinct language families (Yang, 2014) but also because business translation may not fully fit the general translation theory due to its distinct lexical, stylistic, and textual characteristics (Gao, 2018). More attention should be paid to the norms of business translation between English and Chinese.

Every language has its conventions for expressing how people and events relate to one another, which cannot be disregarded if the translation is to be understood (Callow, 1974). Lexical cohesion, as one of five cohesive devices that are reference, substitution, ellipsis, conjunction, and lexical cohesion, means the continuity of text on the level of lexis (Halliday, 1985). In translation, lexical cohesion in a certain type of text assumes an important role in establishing the required effect (Lotfipour-Saedi, 1997). The choice of vocabulary from a shared semantic region substantially makes the text more understandable (Callow, 1974). Translators should generate successful target texts with appropriate cohesive devices (Hu, 1999) because studying lexical cohesion in the translation of a particular field can contribute to the understanding of translation norms.

The translation industry is expected to heavily utilize machine translation in the future (Chan, 2017; Prodanovic et al., 2024). Machine translation enjoys many advantages, including cost-effectiveness, efficiency, and robustness (Qin et al., 2019); it becomes more accurate both in vocabulary and grammar (Groves & Mundt, 2015). In recent years, ChatGPT's translation has gradually become a research hotspot. ChatGPT is a model proficient in diverse Natural Language Processing tasks and encompassing multiple languages, thus essentially functioning as a unified multilingual machine translation model (Jiao et al., 2023). It enhances fluency in the output (Castilho et al., 2023), with over 100 million users in January and accumulating 1 billion visits by February 2023 (Herbold et al., 2023). The model has attracted considerable interest due to its capacity to produce coherent and context-aware texts in translation (Hendy et al., 2023). The fourth generation of GPT is reported to greatly enhance translation performance, reaching a level comparable to that of commercial translation products, even for linguistically distant languages (Jiao et al., 2023). The latest version of ChatGPT, named GPT-4o, was released in May 2024. Scholars should explore the functions and limitations of online digital translation tools to better take advantage of them (Tsai, 2019).

Comparing human translations with AI translations allows us to delve into the strengths and weaknesses of both approaches. Examining cohesion in textual relationships is crucial for both human and machine translation (Menzel et al., 2017). Evaluating human translation and machine translation in English-Chinese coherence not only reveals the differences between the two languages but also explores the

similarities and differences in their specific translation procedures. Discourse analysis of business translation should be done to further explore its translation norms (Ameen & Sherwani, 2023). This study conducts empirical research on the lexical cohesion of English-Chinese business translation by comparing the English texts with two Chinese versions, one by human translators and the other by ChatGPT. Halliday's (1985) theory of cohesion and Toury's (2012) descriptive translation studies are followed to carry out the study. The research questions are as follows:

1) What is the lexical cohesion in English-Chinese business translation?

2) Are there any shifts in lexical cohesion in the English-Chinese business translation by human translators and ChatGPT?

The research seeks to achieve the following research objectives:

1) to describe the lexical cohesion in English-Chinese business translation;

2) to investigate whether there are shifts of lexical cohesion in the English-Chinese business translation by human translators and ChatGPT.

2. Literature Review

2.1 Theoretical Framework

Cohesion comprises grammatical cohesion and lexical cohesion (Halliday, 1985). Lexical cohesion is achieved through word choices of two kinds: reiteration involves repeating a word or using a synonym within a context to refer to the same thing; collocation refers to words that tend to occur in similar lexical contexts (Halliday & Hasan, 1976). Halliday (1985) further sorts lexical cohesion into three categories: repetition, synonymy, and collocation. The repetition refers to the occurrence of the same lexical items. For example, "Algy met a bear. The bear was bulgy", the second "bear" harks back to the first (Halliday, 1985, p. 310). It should be noted that the different morphological shapes of a lexical item also can be taken as repetition, such as dine, dining, diner, and dinner (Halliday, 1985). The synonymy has four subcategories, including synonymy, antonymy, meronymy, and hyponymy, respectively reflecting lexical relations of similarity, opposition, inclusion, and subordination. To be specific, synonymy means two words expressing the same meaning; antonymy refers to lexical items opposite in meaning; meronymy describes the relation of the part and the whole; hyponymy explains the relation of the specific and the general (Halliday, 1985). The collocation stresses a co-occurrence tendency between lexical items. Strong collocation ties cover not only independent words but also fixed phrases and cliches (Halliday, 1985). Collocations are frequently and rather specifically linked to one or more specific registers or functional varieties of the language, and common lexical elements frequently emerge in various collocations depending on the text variety (Halliday, 1985). The types of lexical cohesion are shown in Table 1.

Table 1. Types of lexical cohesion (adapted from Halliday, 1985)

	Types		Examples
Lexical cohesion	Repetition	/	Algy met a bear. The bear was bulgy.
	Synonymy	Synonymy	sound — noise, cavalry — horse
		Antonymy	woke — asleep, sound — silence
		Meronymy	player — team, fountain — garden
		Hyponymy	pine — tree, bed — furniture
	Collocation	/	cold — ice, a stretch of the imagination

Toury's (2012) descriptive translation studies (DTS) explain what occurs during the translation process, as opposed to prescriptive ideas that classify translations as either right or wrong. Translation is a norm-governed behavior (Crisafulli, 2002). In DTS, "textual-linguistic norms govern the selection of linguistic material for the formulation of the target text, or the replacement of the original material" (Toury, 2012, p. 83). Exploring the choice of linguistic elements in the target text, including "lexical items, phrases and stylistic features" contributes to the comprehension of textual-linguistic norms (Munday et al., 2022, p. 154). In addition, DTS encourages in-depth examination of actual texts to understand the translation process and introduces a three-phase methodology for systematic description of translation studies as follows: (1) situating the text within the target culture system, (2) textual analysis of the source text and the target text, and (3) generalizations (Toury, 2012). In the second phase, corresponding segments between the source text and the target text are compared as coupled pairs, which facilitate the recognition of translation shifts. Coupled-pairs method that investigates translational products through a segment-by-segment basis (Hill-Madsen, 2020) reconstructs the decision-making process in translation (Toury, 2012). Combining lexical cohesion with DTS can inspire translation norms of lexical cohesion.

2.2 Related Studies

Scholars have noticed the significance of lexical cohesion in translation. Lotfipour-Saedi (1997) discusses the translation equivalence of lexical cohesion and mentions that lexical cohesion in a text can be analyzed by identifying a central lexical chain based on the semantic relationships between its nodes, the distances between these nodes, and their integration into the text's hierarchy. Pirmoradian and

Dastjerdi (2014) compare the English novel, *The Picture of Dorian Gray*, with its two Persian translations by Tahami and Mashayekhi to examine the methods of cohesive devices used by the two translators. The results demonstrate that the patterns of collocation and reiteration differ in the source text, target texts, and even between the two target texts. This is because some reiterations of the source text in the two target texts are not taken into account by the translators and because there is less variation in the patterns of collocation between the source text and target texts than in the patterns of reiteration. Orang'i and Ndlovu (2021) try to determine whether there are any differences in the usage of lexical cohesion in English-Swahili healthcare texts and characterize the network of lexical chains in those texts. The results show that source texts and target texts are the same in lexical cohesion and compared to their English counterparts, Swahili healthcare texts have a little bit more vocabulary elements. Mahamdeh (2022) looks into cohesive devices in English-Arabic legal translation. The research finds that legal language relies extensively on lexical cohesion of repetition and lexical cohesion is influenced by some language peculiarities and norms. Tarawneh and Al-Momani (2023) probe into translation shifts in the Arabic-English legal translation of lexical repetition. They find that in legal texts English and Arabic employ lexical repetition similarly, the most prevalent shift identified is the partial shift, and the translator uses various translation methods to translate lexical repetition. Sugiarto and Siregar (2023) describe the structure of lexical cohesion in English-Indonesia machine translation and post-editing outputs. They gather data from J.K. Rowling's *Harry Potter and the Order of the Phoenix* and discover that there aren't many discrepancies between the lexical cohesion in the source texts and target texts. These studies cover literary and non-literary texts and prove the significance of lexical cohesion in translation studies among different language pairs. However, research on lexical cohesion in business translation particularly regarding English-Chinese translation and comparison between human translation and ChatGPT's translation is rarely conducted, which makes this study innovative and meaningful.

3. Research Design

3.1 The Corpus for the Research

Parallel corpora offer values in the extraction of translation equivalents (Teubert, 2002), investigation of translation shifts (Munday, 1998), and research into data-driven machine translations (Baker & Saldanha, 2019). In this study, a parallel corpus is built including the excerpts of *Asian Development Bank Annual Report 2021*, *IMF Annual Report 2022*, and *IFC Annual Report 2023* and their two Chinese versions by human translators and ChatGPT, shown in Table 2. According to Flowerdew's (2004) parameters of a specialized corpus, texts that are selected all conform to the parameters: (1) specific purpose for compilation: to gather cohesive devices in the authentic business texts; (2) contextualization: to describe business activities in countries, companies, organizations and individuals with the purpose to disseminate the business information like financial performance, strategic initiatives, plans, potential risks, regulatory requirements, and so forth; (3) genre: reports; (4) type of text/discourse: informative text type; (5) subject matter/topic: business; and (6) variety of language: native English and Chinese translation. The excerpts were the complete six passages randomly selected.

Table 2. The corpus of the study

Constitution	the English source text (ST)	the Chinese target text by human translators (HT)	the Chinese target text by ChatGPT (GT)
Detailed information	1. The passages in *Asian Development Bank Annual Report 2021* include: (1) Partnering with the Private Sector and (2) Expanding Microfinance in Challenging Times. 2. The passages in *IMF Annual Report 2022* include: (1) Economic Surveillance and (2) Lending. 3. The passages in *IFC Annual Report 2023* include: (1) Sustainability and (2) Leveraging AI and Data Science to Drive Better Outcomes.	1. The passages in the Chinese version of *Asian Development Bank Annual Report 2021* translated by Asian Development Bank involve: (1) 与私营部门建立合作伙伴关系 and (2) 扩大小额信贷，助力企业渡过困难期. 2. The passages in the Chinese version of *IMF Annual Report 2022* translated by the International Monetary Fund involve: (1) 经济监督 and (2) 贷款活动. 3. The passages in the Chinese version of *IFC Annual Report 2023* translated by International Finance Corporation involve: (1) 可持续性 and (2) 利用人工智能和数据科学推动更好成果.	GT refers to the whole ST translated by ChatGPT (the version ChatGPT-4o, https://chatgpt.com) that is produced by authors.
Total word numbers	5305	10386	9428

3.2 The Research Procedure

The study conducts qualitative research. Data of lexical cohesion involve lexical items of repetition, synonymy, and collocation according to Halliday's (1985) classification and definition. Lexical items can present cohesive relations of different types simultaneously (Halliday, 1985). To avoid repeatedly counting the actual words, data of this situation are counted into categories with priority sequences of repetition, synonymy, and collocation. Data are collected passage by passage and are made into coupled pairs including items of lexical cohesion in ST, HT, and GT. "The use of semantic and discourse tagging is also becoming more common, but it is still done mainly manually" (Saldanha & O'Brien, 2013, p. 78). Data are identified and collected manually following a four-step procedure shown in Table 3. In the first step, authors identify all items of lexical cohesion in the ST by reading the entire ST. In the second step, authors mark different items of lexical cohesion in the ST and HT by comparatively reading ST and HT. In the third step, the authors mark different items of lexical cohesion in the ST and GT. GT is translated into Chinese by ChatGPT passage by passage. The version of ChatGPT is ChatGPT-4o. GT was produced on June 25, 2024, in Penang. Throughout the translation process, researchers begin each page of the source text in ChatGPT's chat window with the instruction "Translate the business text into Chinese:", which helps signal the specialized nature of the text, ensuring that the translation fits the business context. The authors discovered that ChatGPT sometimes alters its translations when given the same source text repeatedly. Therefore, the translations are based on the first version produced by ChatGPT. Then, the authors mark different items of lexical cohesion in the ST and GT by comparatively reading ST and GT. The fourth step is to classify all lexical items of cohesion in the ST, HT, and GT into three categories according to the types of lexical cohesion and count their word numbers. Authors hold the same standard while they collect and check the data. Qualitative analysis is conducted on the semantic meanings of items of lexical cohesion and their translation shifts in ST, HT, and GT.

Table 3. The research procedure

Steps	Data	Data identification and collection
Step 1	all items of lexical cohesion in ST	Authors identify and mark all items of lexical cohesion in ST by reading the entire ST.
Step 2	different items of lexical cohesion in ST and HT	Authors identify and mark the items of lexical cohesion of translation shifts in HT by comparatively reading ST and HT.
Step 3	different items of lexical cohesion in ST and GT	Authors produce GT. Authors identify and mark the items of lexical cohesion of translation shifts in GT by comparatively reading ST and GT.
Step 4	the classification, count, and analysis of items of lexical cohesion in ST, HT, and GT	The marked items of lexical cohesion in the ST, HT, and GT are classified into different types, counted, and analyzed.

3.3 The Research Limitations

The texts in the corpus do not cover all types of business texts. Involving both written and spoken forms (McEnery & Hardie, 2012; Grygiel, 2015), business discourses include the language used in letters, reports, academic textbooks (Johns, 1980), interviews, negotiations, business meetings, use of electronic media (Grygiel, 2015), as well as conversations of people in business organizations (Boden, 1994). This study excludes phonological and multimodal studies and limits itself to English-written business reports and their Chinese translations. Therefore, the results focusing on written business reports only reveal parts of norms of the cohesion in the English-Chinese translation of overall business texts.

4. Results and Discussions

There are translation shifts of lexical cohesion both in HT and GT with the word numbers of lexical items shown in Table 4. This research adopts purposive sampling. As there are several cases, five examples are provided to describe various types of lexical cohesion and explain the translation shifts in the corpus. The research uses, therefore, the purposive sampling approach (Saldanha & O'Brien, 2013). The examples are selected to reflect all types of lexical cohesion, that occur in the texts, and to describe various kinds of translation procedures achieved by using shifts. All the items of lexical cohesion in the provided examples and the items with translation shifts in the explanations are underlined and bold.

Table 4. The word numbers of items of lexical cohesion with translation shifts

Types of lexical cohesion	HT	GT
Repetition	88	15
Synonymy	98	39
Collocation	166	84
Total word numbers (proportion)	352 (about 3.39%)	138 (about 1.46%)

Example 1

ST: On top of those **programs**, almost a quarter of **ADB**'s 35 nonsovereign **projects** committed in 2021 helped **companies** address **pandemic impacts**. For instance, **ADB** signed **debt financing** of $13.7 million for **Global Health Private Ltd**, which operates **hospitals and clinics** across India under the **Medanta** brand. The **ADB financing** is helping **Medanta** provide 414 **isolation beds** and 281 **ventilators** for its **hospitals and clinics**, train **medical staff** on **COVID-19 infection control**, and maintain existing **health-care services**.

HT: 除上述**计划**外，在**亚行** 2021 年承诺的 35 个非主权**项目**中，近四分之一旨在帮助**企业**应对**疫情影响**。例如，**亚行**与**全球健康有限公司（Global Health Private Ltd）**签署了一项 1,370 万美元的**债务融资协议**。该**公司**以**"曼德塔"（Medanta）**为品牌商标，在印度各地经营**医院和诊所**。**亚行提供的资金**正在帮助**曼德塔**为其**医院和诊所**配置 414 张**隔离病床**和 281 台**呼吸机**，对**医务人员**进行有关**防控新冠病毒感染**的培训，以及维持现有的**医疗卫生服务**。

GT: 除了这些**计划**，**ADB** 在 2021 年承诺的 35 个非主权**项目**中，几乎四分之一帮助**企业**应对**疫情影响**。例如，**ADB** 为 **Global Health Private Ltd** 签署了 1370 万美元的**债务融资**，该**公司**在印度以 **Medanta** 品牌运营**医院和诊所**。**ADB 的融资**帮助 **Medanta** 为其**医院和诊所**提供了 414 张**隔离床位**和 281 台**呼吸机**，培训**医务人员控制 COVID-19 感染**，并维持现有的**医疗服务**。

There are three relations of repetition, four relations of synonymy, and one relation of collocation with all lexical items shown in Table 5. Some shifts occur in the translation of synonymy and hyponymy. For translating synonymy items of "debt financing" and "ADB financing", human translators use "债务融资协议" and "亚行提供的资金", whereas ChatGPT adopts "债务融资" and "ADB 的融资". Different from ChatGPT, human translators add "协议" which means "agreement", the property of "debt financing". It shows that human translators consider the Chinese collocation of "签署协议" to be more common and natural. For translating hyponymy items of "companies" and "Global Health Private Ltd", both human translators and ChatGPT add a lexical item "公司" that means "company". This happens when "Global Health Private Ltd" has an attributive clause and human translators and ChatGPT both render it into an independent clause. It shows that human translators and ChatGPT concur that the English-Chinese translation of syntactic structure requires the addition of lexical cohesion. Human translators and ChatGPT are also different in translating proper nouns of "ADB", "Medanta", and "Global Health Private Ltd" with human translations using Chinese translations "亚行", "曼德塔", and "全球健康有限公司" but ChatGPT's translation keeping the English texts "ADB", "Medanta", and "Global Health Private Ltd".

Table 5. The items of lexical cohesion in Example 1

Lexical cohesion		ST	HT	GT
Repetition	/	(1) two "ADB" (2) two "hospital and clinics" (3) two "Medanta"	(1) two "亚行" (2) two "医院和诊所" (3) "曼德塔(Medanta)" and "曼德塔"	(1) two "ADB" (2) two "医院和诊所" (3) two "Medanta"
Synonymy	Synonymy	(1) "programs" and "projects" (2) "debt financing" and "ADB financing"	(1) "计划" and "项目" (2) "债务融资**协议**" and "亚行提供的资金"	(1) "计划" and "项目" (2) "债务融资" and "ADB 的融资"
	Meronymy	(1) "isolation beds", "ventilator", "medical staff", and "hospitals and clinics"	(1) "隔离病床", "呼吸机", "医务人员", and "医院和诊所"	(1) "隔离床位", "呼吸机", "医务人员", and "医院和诊所"
	Hyponymy	(1) "companies" and "Global Health Private Ltd"	(1) "企业", "全球健康有限公司（Global Health Private Ltd）", and "**公司**"	(1) "企业", "Global Health Private Ltd", and "**公司**"
Collocation	/	(1) "pandemic impacts", "COVID-19 infection control", and "health-care services"	(1) "疫情影响", "防控新冠病毒感染", and "医疗卫生服务"	(1) "疫情影响", "控制 COVID-19 感染", and "医疗服务"

Example 2

ST: **Article IV consultations** consist of a two-way **policy dialogue** between **the IMF** and **country authorities** and cover a range of important **issues**: **fiscal**, **financial**, **foreign exchange**, **monetary**, and **structural**. In **FY 2022**, **the IMF** conducted 126 **Article IV consultations** and 11 **financial system stability assessments** under the **FSAP**.

HT: **第四条磋商**由 **IMF** 与**国家当局**之间的双向**政策对话**构成，涉及**金融**、**财政**、**外汇**、**货币**和**结构性问题**等一系列重要**议题**。**2022 财年**，**IMF** 在 **FSAP** 下开展了 126 次**第四条磋商**和 11 次**金融体系稳定性评估**。

GT: **第四条磋商**包括 **IMF** 与**国家当局**之间的双向**政策对话**，涵盖**财政**、**金融**、**外汇**、**货币**和**结构**等一系列重要**问题**。在 **2022 财年**，**IMF** 进行了 126 次**第四条磋商**，并在**金融部门评估计划（FSAP）**下进行了 11 次**金融系统稳定评估**。

There are two relations of repetition and two relations of collocation with all lexical items shown in Table 6. Some shifts occur in the translation of collocation. For translating collocation items of "fiscal" and "financial", human translators convert their literal meanings into "金融" and "财政", whereas ChatGPT adopts literal translations of "财政" and "金融". Human translators and ChatGPT adopt different translations, but "fiscal" and "financial" are synonyms in the coordination structure, so the translation effects are similar. For translating the collocation item of "structural", human translators use "结构性问题", whereas ChatGPT adopts "结构". Different from ChatGPT's literal translation, human translators add "问题" which means "problem", the property of "structural". It shows that compared to ChatGPT, human translators are more inclined to add attributes to nouns.

Table 6. The items of lexical cohesion in Example 2

Lexical cohesion	ST	HT	GT
Repetition	(1) two "Article IV consultations" (2) two "the IMF"	(1) two "第四条磋商" (2) two "IMF"	(1) two "第四条磋商" (2) two "IMF"
Collocation	(1) "policy dialogue", "country authorities", and "issues" (2) "fiscal", "financial", "foreign exchange", "monetary", "structural", "FY 2022", "financial system stability assessments", and "FSAP"	(1) "政策对话", "国家当局", and "议题" (2) "**金融**", "**财政**", "外汇", "货币", "结构性**问题**", "2022 财年", "金融体系稳定性评估", and "FSAP"	(1) "政策对话", "国家当局", and "问题" (2) "财政", "金融", "外汇", "货币", "结构", "2022 财年", "金融系统稳定评估", and "金融部门评估计划(FSAP)"

Example 3

ST: As **global growth** recovered **during 2021**, **demand** for **IMF financing** gradually shifted from **emergency financing** triggered by urgent, **pandemic-related** balance of **payments** needs to **upper-credit-tranche (UCT)-quality arrangements**, including to deal with **scarring** from the **pandemic**. **Growth** is expected to slow **during 2022**, largely as a consequence of **COVID-19 variants** and the war in Ukraine, which have led to new **lending requests**, including for **emergency financing**.

HT: 随着**全球经济**在 **2021 年**开始复苏，对 **IMF 融资**的**需求**逐渐从与**疫情有关**的**国际收支紧急融资**需求，转向**高信贷档（UCT）标准的安排**，包括用于应对**疫情**造成的**经济创伤**。**全球经济增速**预计将在 **2022 年**放缓，主要是**变异新冠毒株**和乌克兰战争所致。因此，一些国家已经提出新的**贷款请求**，包括**紧急融资请求**。

GT: 随着 **2021 年全球增长**的恢复，**IMF 融资需求**逐渐从因**疫情**紧急**国际收支**需求引发的**紧急融资**转向用于应对**疫情**造成的**创伤**的**上层信贷安排（UCT）质量安排**。预计 **2022 年增长**将放缓，主要原因是 **COVID-19 变种**和乌克兰战争，这些因素导致了新的**贷款请求**，包括**紧急融资**。

There is one relation of repetition, two relations of synonymy, and three relations of collocation with all lexical items shown in Table 7. Some shifts occur in the translation of repetition, synonymy, and collocation. For translating repetition items of two "emergency financing", human translators use "紧急融资" and "紧急融资请求", whereas ChatGPT adopts two "紧急融资". Different from ChatGPT's literal translation that maintains repetition, human translators add "请求" which means "requests" to the translation of the second "emergency financing", which is the purpose of "emergency financing" presented in the same sentence. It shows that human translators prefer adding explanations concerning the functions of nouns compared to ChatGPT. For translating synonymy items of "global growth" and "growth", human translators use "全球经济" and "全球经济增速", whereas ChatGPT adopts "全球增长" and "增长". Different from ChatGPT's literal translation, human translators add "经济" which means "economy" or "economic", serving as the objects of "global growth" and "growth". It shows that human translators tend to add descriptions regarding the behavioral objects of nouns. For translating the collocation items, human translators and ChatGPT add "国际" which means "international" to the translation of "payments" and "经济" which means "economy" or "economic" to the translation of "scarring". It shows that human translators prefer emphasizing the business contexts of lexical items.

Table 7. The items of lexical cohesion in Example 3

Lexical cohesion		ST	HT	GT
Repetition	/	(1) two "emergency financing"	(1) "紧急融资" and "紧急融资**请求**"	(1) two "紧急融资"
Synonymy	Synonymy	(1) "global growth" and "growth" (2) "demand" and "requests"	(1) "**全球经济**" and "**全球经济增速**" (2) "需求" and "请求"	(1) "全球增长" and "增长" (2) "需求" and "请求"
Collocation	/	(1) "during 2021" and "during 2022" (2) "IMF financing", "payments", "upper-credit-tranche (UCT)-quality arrangements", "scarring", and "lending" (3) "pandemic-related", "pandemic", and "COVID-19 variants"	(1) "2021 年" and "2022 年" (2) "IMF 融资", "**国际**收支", "高信贷档(UCT)标准的安排", "**经济**创伤", and "贷款" (3) "疫情有关", "疫情", and "变异新冠毒株"	(1) "2021 年" and "2022 年" (2) "IMF 融资", "国际收支", "上层信贷安排(UCT)质量安排", "创伤", and "贷款" (3) "疫情", "疫情", and "COVID-19 变种"

Example 4

ST: To **advance efforts** in identifying and **managing labor risks**, **IFC's Managing Director and the International Labour Organization (ILO) Director General** agreed on **a Partnership Roadmap** in March 2023. As part of this **roadmap**, **IFC and ILO** are **discussing opportunities** for further **collaboration** and **synergies** for knowledge **sharing** and **exchange**, capacity building, and promoting robust **labor risk management**.

HT: 为了**推进**识别和**管理劳工风险方面**的**工作**，**IFC 董事总经理和国际劳工组织（ILO）总干事**于 2023 年 3 月商定了**一份合作路线图**。作为该**路线图**的一部分，**两机构**正在**讨论**在知识**共享**与**交流**、能力建设以及推动有力的**劳工风险管理等方面**存在哪些进一步**合作**并发挥**协同**效应的**机会**。

GT: 为了**推进**识别和**管理劳工风险**的**工作**，**IFC 总经理和国际劳工组织（ILO）总干事**于 2023 年 3 月同意了**一条合作路线图**。作为该**路线图**的一部分，**IFC 和 ILO** 正在**讨论**进一步**合作**和知识**共享**、能力建设以及促进强有力的**劳工风险管理**的**机会**。

There is one relation of repetition, four relations of synonymy, and one relation of collocation with all lexical items shown in Table 8. Some shifts occur in the translation of repetition, synonymy, meronymy, and collocation. For translating repetition items of "managing labor risks" and "labor risk management", human translators use "管理劳工风险方面" and "劳工风险管理等方面", whereas ChatGPT adopts "管理劳工风险" and "劳工风险管理". Different from ChatGPT's literal translation that keeps the repetition, human translators add "方面" which means "aspect" and "等方面" which means "and other aspects". It shows that human translators favor adding the description of the attributes of nouns. For translating synonymy items of "collaboration" and "synergies", human translators use "合作" and "协同", whereas ChatGPT adopts one "合作". For translating synonymy items of "sharing" and "exchange", human translators use "共享" and "交流", whereas ChatGPT adopts one "共享". Different from human translators' literal translation, ChatGPT uses omission as "collaboration" and "synergies" are synonyms and "sharing" and "exchange" are synonyms in the structure of coordination. It shows that ChatGPT tends to omit the translation of nouns in the coordination structure. For translating meronymy items of "IFC's Managing Director and the International Labour Organization (ILO) Director General" and "IFC and ILO", human translators use "IFC 董事总经理和国际劳工组织（ILO）总干事" and "两机构", whereas ChatGPT adopts "IFC 总经理和国际劳工组织（ILO）总干事" and "IFC 和 ILO". Different from ChatGPT's literal translation, human translators use the conversion method of translating "IFC" and "ILO" in the coordination structure to "两机构" which means "two organizations". It shows that human translators understand the meronymy relation across the two adjacent sentences and consider that the conversion method of nouns in the structure of coordination expresses the similar translation effects of the literal translation method. For translating collocation items of "advance efforts", human translators and ChatGPT both use the conversion translation "推进工作" rather than the literal translation "推进努力". It shows that human translators and ChatGPT have a similar understanding of Chinese collocation and use the conversion method.

Table 8. The items of lexical cohesion in Example 4

Lexical cohesion		ST	HT	GT
Repetition	/	(1) "managing labor risks" and "labor risk management"	(1) "管理劳工风险**方面**" and "劳工风险管理**等方面**"	(1) "管理劳工风险" and "劳工风险管理"
Synonymy	Synonymy	(1) "a Partnership Roadmap" and "roadmap" (2) "collaboration" and "synergies" (3) "sharing" and "exchange"	(1) "一份合作路线图" and "路线图" (2) "合作" and "协同" (3) "共享" and "交流"	(1) "一条合作路线图" and "路线图" (2) "**合作**" (3) "**共享**"
	Meronymy	(1) "IFC's Managing Director and the International Labour Organization (ILO) Director General" and "IFC and ILO"	(1) "IFC 董事总经理和国际劳工组织（ILO）总干事" and "**两机构**"	(1) "IFC 总经理和国际劳工组织（ILO）总干事" and "IFC 和 ILO"
Collocation	/	(1) "advance efforts" and "discussing opportunities"	(1) "推进**工作**" and "讨论机会"	(1) "推进**工作**" and "讨论机会"

Example 5

ST: Since its launch in 2022, **MALENA** has analyzed over 246,000 **internal IFC** and **public documents**—including impact **assessments**, news **articles**, and sustainability **reports**—which stretch back decades and cover over 10,000 projects in 186 countries. One of **MALENA**'s big strengths is **speed**: it reads 19,000 **sentences in a minute**. After scouring the **documents**, **MALENA** generates **dashboards** on **ESG performance**. These **dashboards** can be **a helpful reference source** for **IFC's ESG experts** as they make their **risk assessments**.

HT: **MALENA** 自 2022 年推出以来已经分析了超过 246,000 份 **IFC 内部文件**和**公开文件**（包括影响**评估**、新闻**报道**和可持续性**报告**）。**这些文件**最早可以追溯到几十年前，涵盖了 186 个国家的 10,000 多个项目。**MALENA** 的一大优势是**速度**：**每分钟**可以读取 19,000 个**句子**。**文件**读取完毕后，**MALENA** 会生成有关 **ESG 绩效**的**仪表板**。当 **IFC 的 ESG 专家**进行**风险评估**时，这些**仪表板**可以成为**有用的参考**。

GT: 自 2022 年推出以来，**MALENA** 已分析了超过 246,000 份 **IFC 内部**和**公共文件**，包括影响**评估**、新闻**文章**和可持续性**报告**，**这些文件**涵盖了 186 个国家的 10,000 多个项目，时间跨度数十年。**MALENA** 的一大优势是**速度**：它**每分钟**可以阅读 19,000 **句话**。在仔细审查这些**文档**后，**MALENA** 生成了 **ESG 表现**的**仪表板**。这些**仪表板**可以作为 **IFC 的 ESG 专家**进行**风险评估**时的**有用参考资料**。

There are three relations of repetition, three relations of synonymy, and two relations of collocation with all lexical items shown in Table 9. Some shifts occur in the translation of repetition and collocation. For translating repetition items of two "documents", human translators use three words "文件", "这些文件", and "文件", whereas ChatGPT adopts three words "文件", "这些文件", and "文档". Although "文

件” and “文档” are synonyms with the meaning of “documents” or “papers”, ChatGPT does not keep the repetition. Besides, both human translators and ChatGPT add “这些文件” which means “these documents”. This happens when the first word “documents” has an attributive clause and human translators and ChatGPT render it into an independent clause. It indicates that both human translators and ChatGPT agree that translating syntactic structures from English to Chinese necessitates the addition of lexical cohesion. For translating the collocation item of “a reference source”, human translators use “参考”, whereas ChatGPT adopts “参考资料”. Different from ChatGPT’s literal translation, human translators omit the translation of “资料”. It shows that human translators stress the function of the “source” of reference rather than its material property.

Table 9. The items of lexical cohesion in Example 5

Lexical cohesion		ST	HT	GT
Repetition	/	(1) three “MALENA” (2) two “documents” (3) two “dashboards”	(1) three “MALENA” (2) “文件”, “**这些文件**”, and “文件” (3) two “仪表盘”	(1) three “MALENA” (2) “文件”, “**这些文件**”, and “**文档**” (3) two “仪表盘”
Synonymy	Antonymy	(1) “internal IFC” and “public”	(1) “IFC 内部” and “公开”	(1) “IFC 内部” and “公共”
	Meronymy	(1) “sentences” and “documents”	(1) “句子” and “文件”	(1) “句话” and “文档”
	Hyponymy	(1) “documents”, “assessments”, “articles”, and “reports”	(1) “文件”, “评估”, “报道”, and “报告”	(1) “文件”, “评估”, “文章”, and “报告”
Collocation	/	(1) “speed” and “in a minute” (2) “ESG performance”, “a helpful reference source”, “IFC’s ESG experts”, and “risk assessments”	(1) “速度” and “每分钟” (2) “ESG 绩效”, “有用的**参考**”, “IFC 的 ESG 专家”, and “风险评估”	(1) “速度” and “每分钟” (2) “ESG 表现”, “有用参考资料”, “IFC 的 ESG 专家”, and “风险评估”

In brief, all types of lexical cohesion occur in ST, HT, and GT. Although the specialized texts are business texts, cohesive relations of other specialized types exist, such as cohesive relations involving medical items in Example 1, political items in Example 2, and technological items in Example 5. Both human translators and ChatGPT make a small proportion of translation shifts of lexical cohesion overall, with the most shifts in collocation, followed by synonymy, and the least shifts in repetition. However, in terms of the proportion of translation shift word count to the total word count, human translators handle translation shifts in lexical cohesion more than twice as much as ChatGPT. What’s more, for making specific translation shifts, both human translators and ChatGPT can adopt the methods of literal translation, addition, omission, and conversion in translating lexical cohesion depending on different situations. The literal translation is to use the common meanings of words; addition pertains to adding information; omission refers to reducing the information; conversion concerns paraphrasing the information. In addition to the mostly used method of literal translation, addition is often adopted both by human translators and ChatGPT in different situations. This happens especially when nouns have relatively long attributive clauses and human translators and ChatGPT attempt to translate them into independent clauses. Compared to ChatGPT, human translators prefer adding words to describe the properties and attributes of nouns, which may be one of the reasons that the length of HT is longer than that of GT. The omission is applied sometimes when lexical items suggest unimportant information and occur in parallel structures. Conversion is employed occasionally without altering the cohesive relations. When translating proper nouns, human translators often use Chinese translations, while ChatGPT typically opts to retain English texts.

5. Conclusion

This study explores the English-Chinese translation of lexical cohesion in business texts by human translators and ChatGPT’s translation. The results indicate that both human translators and ChatGPT keep three types of lexical cohesion, which are repetition, synonymy, and collocation, to a large extent, and the translation methods of literal translation, addition, omission, and conversion are employed. Compared to ChatGPT, human translators make more translation shifts than ChatGPT. The implications of this study involve English-Chinese business translation. It offers effective insights into machine and human translation, concerning translation shifts and norms of cohesion in business translation and sheds light on the need for translators’ post-editing activity for lexical cohesion translation.

In future research, translation quality of translation shifts of lexical cohesion between human translation and ChatGPT’s translation should be conducted on English-Chinese business translation. Besides, more text materials such as letters, textbooks, and everyday conversations should be covered to enrich the corpus of research in this field and obtain more generalizable results. Since this research is done on business texts, the authors recommend that other research be done on the translation of different specialized texts to have general conclusions on lexical cohesion translation.

Acknowledgments

Not applicable.

Authors contributions

Na Tang and Mohamed Abdou Moindjie were responsible for study design and revising. Na Tang was responsible for data collection. Na Tang drafted the manuscript and Mohamed Abdou Moindjie revised it. All authors read and approved the final manuscript.

Funding

Not applicable.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Sciedu Press.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

Ameen, J. A. M., & Sherwani, K. A. (2023). Semantic challenges of business translation with reference to English and Arabic: A product-oriented approach. *Journal of Language Studies*, *6*(3), 1-13. https://doi.org/10.25130/jls.6.3.2.1

Baker, M., & Saldanha, G. (2019). *Routledge encyclopedia of translation studies* (3rd ed.). Routledge. https://doi.org/10.4324/9781315678627

Boden, D. (1994). *The business of talk: Organizations in action*. Polity Press.

Callow, K. (1974). *Discourse considerations in Translating the Word of God*. Zondervan Publishing House.

Castilho, S., Mallon, C., Meister, R., & Yue, S. (2023, June 12-15). *Do online machine translation systems care for context? What about a GPT model?* [Paper presentation]. The 24th Annual Conference of the European Association for Machine Translation, Tampere, Finland. Retreived from https://aclanthology.org/2023.eamt-1.39

Chan, A. L. J. (2017). Chinese translation market. In C. Shei & Z. M. Gao (Eds.), *The Routledge handbook of Chinese translation* (pp. 257-271). Routledge. https://doi.org/10.4324/9781315675725-15

Chiper, S. (2002). Business translation. *Perspectives: Studies in Translation Theory and Practice*, *10*(3), 215-233. https://doi.org/10.1080/0907676X.2002.9961446

Crisafulli, E. (2002). The quest for an eclectic methodology of translation description. In T. Hermans (Ed.), *Crosscultural transgressions* (pp. 26-43). St. Jerome Publishing.

Flowerdew, L. (2004). The argument for using English specialized corpora to understand academic and professional Settings. In U. Connor & T. Upton (Eds), *Discourse in the professions: Perspectives from corpus linguistics* (pp. 11-33). John Benjamins Publishing. https://doi.org/10.1075/scl.16.02flo

Gao, L. (2018). A study on the application of functional equivalence to business English E-C Translation. *Theory and Practice in Language Studies*, *8*(7), 759-765. https://doi.org/10.17507/tpls.0807.06

Gotti, M., & Šarčević, S. (Eds.). (2006). *Insights into specialized translation*. Peter Lang. https://doi.org/10.3726/978-3-0351-0411-0

Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes*, *37*, 112-121. https://doi.org/10.1016/j.esp.2014.09.001

Grygiel, M. (2015). Business English from a linguistic perspective. *English for Specific Purposes World*, *16*(1), 1-12.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. Edward Arnold.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.

Hendy, A., Abdelrehim, M., Sharaf, A., Raunak, V., Gabr, M., Matsushita, H., … Awadalla, H. H. (2023). How good are GPT models at machine translation? A comprehensive evaluation. *arXiv preprint*. https://doi.org/10.48550/arXiv.2302.09210

Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z., & Trautsch, A. (2023). A large-scale comparison of human-written versus ChatGPT-generated essays. *Scientific Reports*, *13*(1), 1-11. https://doi.org/10.1038/s41598-023-45644-9

Hill-Madsen, A. (2020). SFL and Descriptive Translation Studies: Systemic-functional grammar as a framework for the analysis of shifts in translation. *Globe: A Journal of Language, Culture and Communication*, *10*, 143-169. https://doi.org/10.5278/ojs.globe.v10i.5883

Hu, H. C. (1999). Cohesion and coherence in translation theory and pedagogy. *Word*, *50*(1), 33-46. https://doi.org/10.1080/00437956.1999.11432482

Jiao, W., Wang, W., Huang, J. T., Wang, X., Shi, S., & Tu, Z. P. (2023). Is ChatGPT a good translator? Yes with GPT-4 as the engine. *arXiv preprint*. https://doi.org/10.48550/arXiv.2301.08745

Johns, A. (1980). Cohesion in written business discourse: Some contrasts. *The ESP Journal*, *1*(1), 35-43. https://doi.org/10.1016/0272-2380(80)90008-6

Lotfipour-Saedi, K. (1997). Lexical cohesion and translation equivalence. *Meta*, *42*(1), 185-192. https://doi.org/10.7202/004014ar

Mahamdeh, A. A. H. (2022). Translation of cohesive devices in selected legal text types from English into Arabic. *International Journal of Linguistics, Literature and Translation*, *5*(4), 1-11. https://doi.org/10.32996/ijllt.2022.5.4.1

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press. https://doi.org/10.1017/CBO9780511981395

Menzel, K., Lapshinova-Koltunski, E., & Kunz, K. (2017). Cohesion and coherence in multilingual contexts. In K. Menzel, E. Lapshinova-Koltunski, & K. Kunz (Eds.), *New perspectives on cohesion and coherence* (pp. 1-10). Language Science Press.

Munday, J. (1998). A computer-assisted approach to the analysis of translation shifts. *Meta*, *43*(4), 542-556. https://doi.org/10.7202/003680ar

Munday, J., Pinto, S. R., & Blakesley, J. (2022). *Introducing translation studies: Theories and applications* (5th ed.). Routledge. https://doi.org/10.4324/9780429352461

Orang'i, D. O., & Ndlovu, M. V. (2021). Lexical cohesion in the translation of English-Swahili health care texts. *Southern African Linguistics and Applied Language Studies*, *39*(3), 269-281. https://doi.org/10.2989/16073614.2021.1942096

Pirmoradian, E., & Dastjerdi, H. V. (2014). A comparative study of the Picture of Dorian Gray and its two Persian translations in terms of cohesive devices. *Mediterranean Journal of Social Sciences*, *5*(20), 2373-2380. https://doi.org/10.5901/mjss.2014.v5n20p2373

Prodanovic, M., Pukienė, A., & Brazauskaitė-Zubavičienė, I. (2024). On challenges in translation: Contrasting human and AI-supported practices. *Revista Publicando*, *11*(43), 46-57. https://doi.org/10.51528/rp.vol11.id2420

Qin, Y., Zhang, J., & Lu, X. (2019, March). *The gap between NMT and professional translation from the perspective of discourse* [Paper presentation]. the 2019 3rd International Conference on Innovation in Artificial Intelligence, New York, United States. https://doi.org/10.1145/3319921.3319936

Saldanha, G., & O'Brien, S. (2013). *Research methodologies in translation studies*. St Jerome Publishing. https://doi.org/10.4324/9781315760100

Sugiarto, B. R., & Siregar, B. U. (2023). Lexical cohesion in English-Indonesia machine translation output: The realization of manual post-editing. *Journal of Applied Linguistics and Literacy*, *7*(1), 174-184. https://doi.org/10.25157/jall.v7i1.9862

Tarawneh, R. T., & Al-Momani, I. M. (2023). Contrastive analysis of translation shifts in lexical repetition in Arabic-English legal translations. *World Journal of English Language*, *13*(1), 69-76. https://doi.org/10.5430/wjel.v13n1p69

Teubert, W. (2002). The role of parallel corpora in translation and multilingual lexicography. In B. Altenberg & S. Granger (Eds.), *Lexis in contrast: Corpus-based approaches* (pp. 189-214). John Benjamins Publishing. https://doi.org/10.1075/scl.7.14teu

Toury, G. (2012). *Descriptive translation studies and beyond* (2nd ed.). John Benjamins Publishing. https://doi.org/10.1075/btl.100

Tsai, S. (2019). Using google translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning*, *32*(5-6), 510-526. https://doi.org/10.1080/09588221.2018.1527361

Yang, J. S. (2014). A contrastive study of cohesion in English and Chinese. *International Journal of English Linguistics*, *4*(6), 118-123. https://doi.org/10.5539/ijel.v4n6p118