PIONEER
ACADEMIC PUBLISHING LIMITED

Journal of Linguistics and
Communication Studies
ISSN 2958-0412
www.pioneerpublisher.com/jlcs
Volume 3 Number 3 September 2024

Faithfulness Fails to Fix it All: A Corpus-Driven Investigation into English Translation Issues in Abstracts of Chinese Civil Engineering Research Articles

Yujun Wang[1]

[1] School of Foreign Studies, Tongji University, Shanghai, China
Correspondence: Yujun Wang, School of Foreign Studies, Tongji University, Shanghai, China.

doi:10.56397/JLCS.2024.09.03

Abstract

The primary concern in translating research article abstracts into English is to ensure accuracy and efficiency. This study established two corpora: CCEJAC (Chinese Civil Engineering Journal Abstract Corpus) and ICEJAC (International Civil Engineering Journal Abstract Corpus). The paper applied Hyland's five-move model and established ten sub-corpora for comparative studies of equivalent moves in CCEJAC and ICEJAC. It was observed that CCEJAC and ICEJAC differ in both lexical and syntactic levels. Specifically, differences were found in the use of general academic vocabulary and bundles, as well as in sentence length and focus. It can be concluded that translation versions in CCEJAC prioritize preserving the literal meaning and the original sentence structures, which may end in undesirable outcomes. These contrasts are the main factors contributing to English translation issues. To address these, this study thoroughly investigated the English translation issues in each move in CCEJAC. The aim is to refine English versions of abstracts to be more concise and practical, in a bid to further facilitate academic communication between Chinese and international scholars in the field of civil engineering.

Keywords: civil engineering abstracts, English translation, corpus-driven, move analysis

1. Introduction

Abstracts serve as concise summaries of academic papers. Moreover, English abstracts have become a prominent means for international academia to comprehend Chinese scholars' core research achievements. Scholars in China have expanded the research on abstract translation across various disciplinary domains. Zhang (2011) emphasized the expository and textual functions in the English translation of sports paper abstracts. Chen and Wang (2018) explored the Chinese-English differences in abstracts of humanities and social sciences research articles from the perspective of translation rhetoric, summarizing corresponding strategies and discourse patterns. In medical paper abstracts, Li and Fan (2014) and Liu (2015) highlighted the characteristics and translation strategies from the perspectives of explicitation and aesthetics, respectively. Xia et al. (2022) took

medical paper abstracts as an example to identify issues in the machine translation of scientific paper abstracts, such as omissions, mistranslations, and mishandling of complex sentences, proposing a human-machine collaborative approach to enhance the translation quality. From the above studies, it is evident that abstract translation in various disciplines has received significant attention in terms of its communicative purposes. However, currently, there are relatively limited systematic translation studies on abstract translation in civil engineering. Furthermore, due to the natural science attributes of civil engineering abstracts, the translation strategies of other disciplines may not be fully applicable to this field. This gap underscores the necessity for targeted research to develop effective translation strategies tailored to civil engineering.

The existing research on English abstracts in civil engineering encompasses both linguistic and translation studies. Yi (2008) analyzed abstracts in international civil engineering journals based on Halliday's systemic functional linguistics. Huangfu et al. (2012) empirically studied English abstracts' language features and genres in Chinese civil engineering research articles. Sun (2015) analyzed specific cases to find the Chinese civil engineering abstracts' translation method and the English versions' linguistic characteristics. Xu and Chen (2019) applied the IMRD move analysis method to compare the language features of abstracts in civil engineering papers written by Chinese scholars and native English speakers, and proposed improvements to the writing forms of the former.

In brief, academic research in civil engineering abstract translation indicates a notable absence of comparative studies on Chinese and international journal papers. Additionally, there is a need for corpus-driven investigations and strategies to resolve prevalent issues in existing translations. Moreover, current studies mainly concentrate on elements such as tense, voice, and sentence patterns, often overlooking the genre-specific characteristics of abstracts. Consequently, they do not adequately address the communicative needs of English abstracts for civil engineering papers. As stated by Song et al. (2020: 108), "Scientific paper abstracts have a distinct purpose and structural function, offering a highly summarized overview of the paper and serving as a crucial reference for readers to access and select papers." Given its typical nature as a scientific paper abstract, the English abstract of civil engineering papers holds significant importance in presenting the research accomplishments of Chinese scholars and fostering international academic exchange.

2. Materials and Methods

This study adopted Hyland's (2000) five-move framework for abstract analysis: Introduction-Purpose-Method-Product-Conclusi -on. The research objects were sourced from: 50 English abstracts from Chinese civil engineering journals indexed by EI Compendex (e.g., *Rock and Soil Mechanics*, *Chinese Journal of Geotechnical Engineering*, *Chinese Journal of Engineering*, etc.), and also 50 abstracts from high-level SCI international civil engineering journals (e.g., *Tunnelling and Underground Space Technology*, *Structures*, *Composites Part B: Engineering*, etc.) They were selected using non-probability sampling. These English abstracts were applied to establish two small-scale corpora: Chinese Civil Engineering Journal Abstract Corpus (named CCEJAC) and International Civil Engineering Journal Abstract Corpus (named ICEJAC). After text cleaning and conversion to .txt format, the author created the following ten sub-corpora according to move analysis: CCEJAC-Introduction, CCEJAC-Purpose, CCEJAC-Method, CCEJAC-Product, CCEJAC-Conclusion, ICEJAC-Introduction, ICEJAC-Purpose, ICEJAC-Method, ICEJAC-Product, and ICEJAC-Conclusion. Text analysis tools, including Voyant, AntConc 3.5.9, and WordSmith 8.0, were employed to investigate lexical and syntactic differences between CCEJAC and ICEJAC.

3. Results

3.1 Contrast of Lexical Features Between CCEJAC and ICEJAC

3.1.1 Differences in the Use of Academic Words

Nation (2013) distinguished between technical words (the vocabulary used in specific academic disciplines) and academic words (words that possess transferability across various academic disciplines). The Academic Vocabulary List (AVL), compiled by Gardner and Davies (2014) based on the academic sub-corpus of COCA, has a larger corpus size compared to the Academic Word List (AWL) by Coxhead (2000). According to Wang & Liu (2024: 91), "AVL's vocabulary coverage in academic texts in both BNC and COCA corpora is higher than AWL."

Furthermore, the AVL's selection criteria for academic vocabulary are more rigorous, excluding general high-frequency vocabulary and technical terms (Gardner & Davies, 2014). It contains commonly used academic words across various disciplines. Hence, this paper selected AVL as the reference for the academic word comparison analysis.

Visual text analysis tool Voyant was used to generate word clouds for CCEJAC and ICEJAC for an intuitive comparison, as shown in Figures 1 and 2.

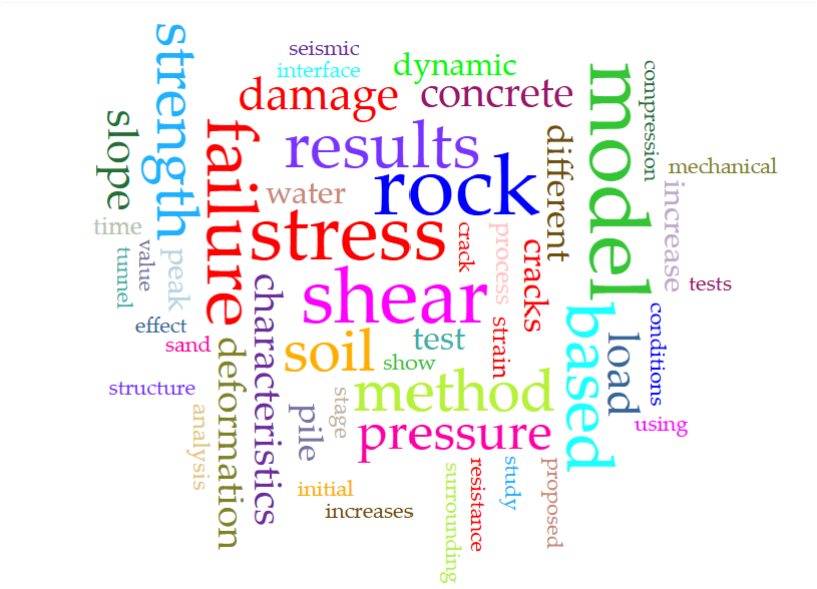

Figure 1. Word Cloud of CCEJAC Generated by Voyant

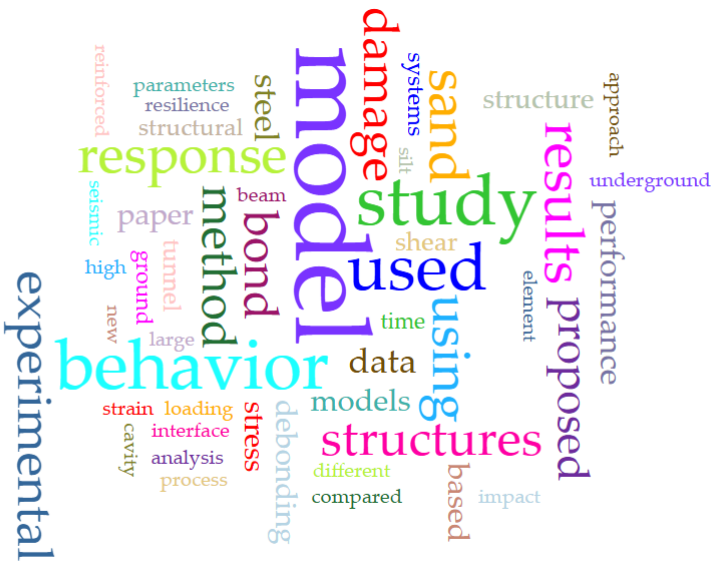

Figure 2. Word Cloud of ICEJAC Generated by Voyant

Figures 1 and 2 visually illustrate the most used vocabulary in the two corpora. The distributions of high-frequency words in CCEJAC and ICEJAC display differences and overlaps. The prevalence of technical words is more pronounced in CCEJAC, whereas ICEJAC demonstrates a contrasting trend, indicating the likelihood of a higher proportion of academic words within its text. In CCEJAC, the top ten frequently used academic words include "model" (86 hits), "results" (71 hits), "based" (64 hits), "method" (56 hits), "characteristics" (44

hits), "different" (42 hits), "increase" (37 hits), "conditions" (32 hits), "tests" (25 hits), and "proposed" (23 hits). Meanwhile, in ICEJAC, the top ten commonly used academic words are "model" (63 hits), "study" (40 hits), "behavior" (39 hits), "used" (35 hits), "results" (32 hits), "structures" (29 hits), "using" (29 hits), "experimental" (28 hits), "method" (26 hits), and "proposed" (25 hits). A comparative analysis revealed that "model", "results", "method", and "proposed" are academic words found in both corpora with the highest frequency, all of which are contained in the AVL and should be given priority in the translation process.

Besides, CCEJAC and ICEJAC vary in their choice of words to convey the same meanings. Using the text analysis tool AntConc 3.5.9, word lists were generated for the ten sub-corpora of CCEJAC and ICEJAC, with common high-frequency words (e.g., articles, prepositions, conjunctions, etc.) and technical words excluded. Upon a closer examination of the remaining academic words, a comparison revealed that the Latin-derived word "utilize" appeared in CCEJAC-Introduction, which was not found in the AVL. As opposed to the former, across various moves in ICEJAC, the preference leaned towards the Old English-derived term "apply", as recorded in the AVL, thus reducing the readers' cognitive load. Furthermore, a search in the Corpus of Contemporary American English (COCA) indicated that the frequency of "apply" was 82.90 occurrences per million in the academic Sci/Tech register. At the same time, "utilize" appeared at a rate of 17.53 per million in the same register, highlighting the greater prevalence of "apply" in academic discourse. Additionally, it was noted that in CCEJAC, the term "way" was used to denote methodology, while in ICEJAC, the more specific and explicit terms "approach", "technique", and "method", as cataloged in the AVL, were commonly employed to represent research methodologies. The above phenomenon mirrors CCEJAC's tendency to use more general terms and ICEJAC's preference for academic words that convey a more precise and specific meaning.

3.1.2 Differences in the Use of General Academic Bundles

Bundles, defined as "frequently recurrent strings of uninterrupted word-forms" (Hyland, 2008: 5), are crucial in academic discourse. Hyland (2008) first introduced research-oriented, text-oriented, and participant-oriented bundles based on their functions. Like academic words, these bundles are extensively employed in the general academic register to convey precise research information, construct logical arguments, and ensure reader comprehension. They help enhance academic discourse's professional level, coherence, and communicative efficiency.

This study analyzed bundles in ten sub-corpora from CCEJAC and ICEJAC by using the text analysis tool AntConc 3.5.9. The Clusters/N-Grams function was applied to identify the appeared bundles, with N-Gram Size set from 2 to 5 (the most commonly seen size). Subsequently, the KWIC function was used to manually verify their contextual co-occurrence and extract semantically complete bundles for further analysis. Following Hyland's (2008) classification of academic bundles by function, research-oriented, text-oriented, and participant-oriented bundles were classified in both CCEJAC and ICEJAC. Given the objective nature of civil engineering research articles, the majority of the bundles fall into the categories of research-oriented and text-oriented ones. Therefore, the analysis primarily focused on these two types of bundles. Upon manual comparison of bundles with similar functions, it was observed that CCEJAC had relatively limited variety and more general meanings in the use of bundles, whereas ICEJAC demonstrated greater diversity and more specific meanings in the use of bundles.

For instance, in the research-oriented bundles, when introducing the research background and prior studies, common bundles in CCEJAC-Introduction involve "the application of" and "the development of", while in ICEJAC-Introduction, more specific bundles can be seen, such as "initial attempts to", "the specific problem of", and "the theoretical analysis". Besides, when describing concepts, the bundle "as an important way of" appears in CCEJAC-Introduction, which, in contrast to the bundle "give good predictions of" in ICEJAC-Introduction, is more general and may not be efficient in reflecting the research's proficiency.

On top of that, in terms of text-oriented bundles, CCEJAC is comparatively lacking in transitional bundles in each move, primarily embodying "based on", "in addition", "according to", "compared with", "compared to", and "in terms of". In stark contrast, ICEJAC showcases a

diverse array of transitional bundles throughout each move, encompassing "with regard to", "coupled with", "compared to", "in addition", "followed by", "when it comes to", "for this purpose", "with respect to", "in accordance with", "in contrast", "along with", "on the basis of", "on the other hand", "in the case of", and more. What's more, CCEJAC is more likely to use abstract or neutral text-oriented bundles, such as "as a result of", "be related to", "caused by", "associated with", "account for", and "provide a reference for" to express cause and effect relationships. By contrast, ICEJAC employs more specific and explicit text-oriented bundles to serve the same function, including expressions like "impinge on", "be dependent on", "be valid with", "be determined by", "provide sufficient evidence to", and "be considerably influenced by".

3.2 Contrast of Syntactic Features Between CCEJAC and ICEJAC

3.2.1 Differences in the Sentence Length

Average sentence length is a crucial metric for evaluating syntactic complexity. This study applied the WordList feature of the text analysis tool WordSmith 8.0 to measure the average sentence length for each sub-corpus of CCEJAC and ICEJAC, as illustrated in Table 1.

Table 1. Average Sentence Length in Each Sub-corpus of CCEJAC and ICEJAC

	Introduction	**Purpose**	**Method**	**Product**	**Conclusion**
CCEJAC	26.65	24.76	26.37	26.76	28.18
ICEJAC	22.63	22.65	19.21	23.87	22.94

"When the average sentence length exceeds 25 words, the text becomes obscure and may lead to misinterpretation (He et al., 2008: 406)." Excessively long sentences are detrimental to conveying the central ideas. As shown in Table 1, the average sentence lengths of sub-corpora in CCEJAC consistently exceed those of ICEJAC, with four sub-corpora in CCEJAC having average sentence lengths surpassing 25 words. However, those in ICEJAC all fall within 25 words. Particularly in the experiment design and procedural sections, CCEJAC-Method and ICEJAC-Method show a significant difference in average sentence length, with a discrepancy of approximately 7.16 words. The results reveal a trend in the CCEJAC favoring longer sentences over those in the ICEJAC. Further manual analysis confirmed the prevalence of excessively long subject phrases and complex compound subject clauses in CCEJAC. This phenomenon, to some extent, diminishes the readability of the text. However, the ICEJAC shows a stark contrast. Excessively long subject phrases are rarely seen, and simple subject clauses are prevalent, except for the occasional compound subject clauses.

In an attempt to explore reasons behind excessively long sentences in CCEJAC, this research employed the Concordance Plot feature of the text analysis tool AntConc 3.5.9 to search for common coordinating conjunctions (e.g., and, but, and or), as well as subordinating conjunctions (e.g., that, which, as, because, etc.). As the number of sentences varies between each move in both CCEJAC and ICEJAC, the character data measured by the Chars function was not factored into the analysis. Instead, this study focused on the frequency of specified word occurrence in the text, as measured by the Hits function. This analysis aimed to determine whether there was an overuse of these conjunctions in CCEJAC, which was thought to be a fundamental cause of the extended sentence length.

Figure 3. Distribution of "and" in CCEJAC's Moves as Generated by AntConc

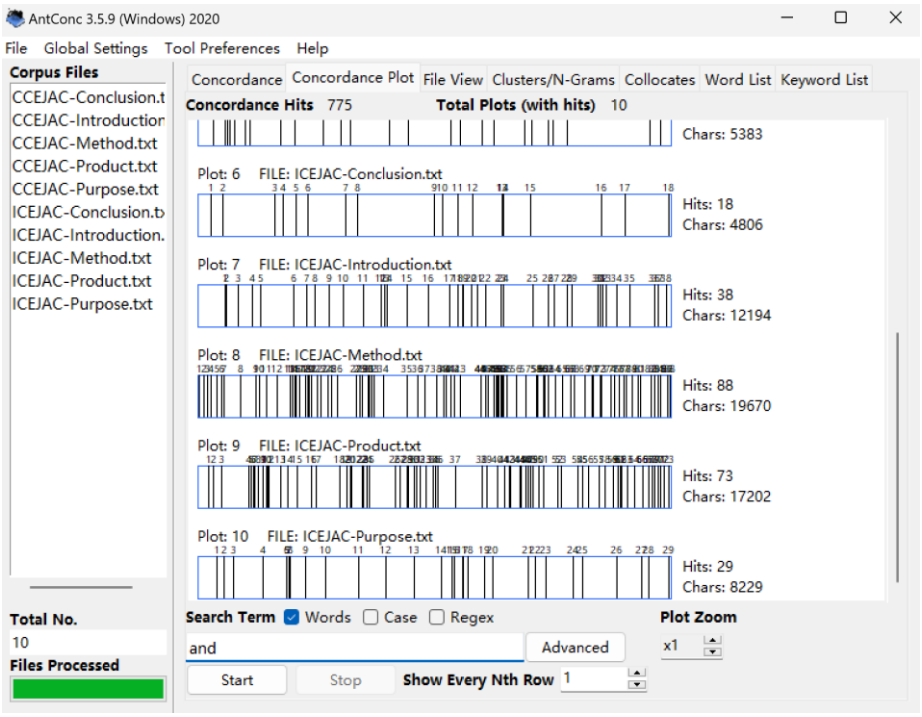

Figure 4. Distribution of "and" in ICEJAC's Moves as Generated by AntConc

Figures 3 and 4 depict the distribution of the coordinating conjunction "and" within the sub-corpora of CCEJAC and ICEJAC, respectively. The two figures highlight the notably high frequency of "and" in CCEJAC-Method and CCEJAC-Product. Furthermore, other moves of CCEJAC also bear a substantial number of compound sentences. Take a sentence from CCEJAC-Introduction for example: "Fly ash is the main waste of thermal power generation, which is widely available in western China, subject to transportation, high disposal costs, and serious desertification problems in the western Ningxia and Gansu regions, and based on the relationship between the supply and demand of the two to carry out research on the application of fly ash-modified materials for sand fixation." This 61-word

sentence lacks text-oriented bundles for coherence, making it hard to convey the dense information through its complex structure. Its organization and structure are constrained by the implicit features of the original Chinese sentence. Hence, it results in a series of components with obscure logical relationships, reducing communication efficiency and posing comprehension challenges.

Meanwhile, through the Concordance Plot, it was found that the CCEJAC-Product contained the highest number of non-restrictive and restrictive subordinate clauses introduced by "which", totaling 16 sentences. Additionally, the search revealed a significant occurrence of noun clauses introduced by "that" in the CCEJAC-Product, amounting to 60 sentences. These findings indicated that the frequent uses of subordinate clauses and noun clauses are primary factors contributing to lengthy sentences. These phenomena may lead to confusion in logical connections between sentences, making it challenging for readers to discern the main points and extract core information. For instance, consider the following sentence from the CCEJAC-Product: "The silicon-oxygen bond produced by the hydrolysis and condensation of siloxane in the cement-based interfacial agent is adsorbed on the hydrophobic layer formed at the rock-concrete interface, which inhibits the expansion of the interface water-rich zone." This 36-word sentence contains extensive information, multiple subordinate clauses, and modifiers, lacking appropriate bundles to separate it into smaller segments. Therefore, it becomes lengthy and intricate, devoid of clear points of emphasis.

3.2.2 Differences in the Sentence Focus

"English sentences adhere to the end-focus principle, and there is a tendency to place the most complex elements at the end to maintain balance in the sentence" (Hu & Leng, 2023: 25). The CCEJAC corpus frequently uses the passive voice and tends to position the predicate at the end of sentences. This extends the dependency distance between the subject and predicate, often resulting in a reduction in the emphasis on critical information. The ICEJAC corpus bears a difference. It involves active and passive voice sentences, with a notable absence of placing verbs at the sentence's end. Therefore, sentences in ICEJAC have enhanced clarity and balance, enabling readers to identify and understand the main points more easily. Take the sentence in CCEJAC-Introduction as an example: "With the continuous development of urban underground space in China, safety problems between urban underground pipelines and underground engineering construction that are in active service are constantly emerging." Due to the lengthy subject and a string of post-modifiers, "the dependency distance between components increases, thereby heightening text complexity" (Zhang et al., 2024: 1119) and causing a shift in the sentence's focus. Similarly, in ICEJAC-Introduction, a sentence serves the same function of stressing the research's significance by describing existing problems. It is presented as: "The model assumes that the speed of sound is infinite which leads to challenges related to solving the equation for pressure". In this sentence, the subject "the model" and the predicate "assumes" are closely linked, and the core information is placed at the end, achieving a more efficient and direct transmission of information.

4. Discussion

The differences between CCEJAC and ICEJAC reveal two crucial challenges in the English translation of civil engineering academic paper abstracts: lexical selection and sentence structure. Improper use of academic words and general academic bundles may lead to semantic misinterpretation and abstract expressions. The absence of general academic bundles may cause semantic incoherence and unclear logic. Furthermore, lengthy sentences primarily manifest in the excessive use of coordination, subordinate clauses, and noun clauses. The displacement of sentence focus is mainly reflected in inappropriate word order, excessive dependency distance, and overuse of the passive voice, all of which can obscure critical information and increase cognitive load. These issues often occur due to the constraints imposed by the sentence structure of the original Chinese text and the strict adherence to literal meanings in the translations. Considering this, it is essential to recognize the distinctions between Chinese and English and adopt suitable translation strategies to address the issues. The goal is to accurately convey academic information and effectively fulfill the communicative function of each move.

The following are typical cases of the original sentences and their translation versions from each move of CCEJAC, and they contain specific issues that could be further refined in the

translation process.

4.1 CCEJAC-Introduction

The Introduction move of civil engineering research articles typically provides an overview of unresolved issues within the field, a review of previous research findings, and identification of research gaps and deficiencies. Its purpose is to introduce the article's topic and underscore the research's significance. The translation version of case 1 is from CCEJAC-Introduction:

Case 1 寒区隧道**普遍受到**冻融作用**影响**，易发生沿着混凝土衬砌结构与围岩界面损伤与脱黏现象，严重**影响**隧道工程施工质量与安全运营。

Translation Version: Tunnels in cold region**s are generally affected by** freeze-thaw action, which is prone to damage and debonding along the interface between concrete lining structure and surrounding rock, which seriously **affects** the construction quality and safe of tunnel engineering operation.

First and foremost, the word "affect" appears twice, which does not align with the norms of English writing. Furthermore, regarding expressions of cause and effect relationships, the neutral meaning of "affect" in the bundle "be generally affected by" should be reconsidered and possibly changed to a more specific and precise term. For instance, "be generally susceptible to" or "be generally subject to" will better convey the detrimental impact of freeze-thaw action, which is more in line with the original intent. Given the sentence structure, the translation version is overly lengthy, with two consecutive subordinate clauses introduced by "which". There could be potential confusion due to the unclear reference of "which". It is suggested that the final subordinate clause be split into a shorter independent sentence to enhance clarity. Besides, it is recommended to add the text-oriented bundle "in light of this" at the beginning of this independent sentence and replace "which" with "it", in a bid to further explain the implicit logical meaning in the original sentence.

4.2 CCEJAC-Purpose

In civil engineering research articles, the Purpose move serves as a guide to the research subject, scope, motivation, and core objectives. It typically includes the proposed theory or hypothesis to be validated, improvements to existing technologies, strategies for specific engineering challenges, etc. The primary communicative purpose of this move is to build a robust foundation for subsequent discussions. Specifically, it aims to render readers a comprehensive and lucid framework of the study, enabling them to grasp its significance, objectives, and expected results. The following is an example of a translation version taken from CCEJAC-Purpose:

Case 2 本文以弱膨胀土为试验材料，开展了宽广吸力范围内非饱和土的持水和强度特性试验研究，并提出适用于宽广吸力范围的非饱和土强度模型。

Translation Version: In this paper, experimental **studies** on the water retention and strength properties of unsaturated weak expansive soil in a wide suction range **were carried out**, and the strength model of unsaturated soil for wide suction range was proposed.

The translation of Case 2 displays a considerable dependency distance of 17 words between the subject "studies" and the predicate "were carried out". This prompts readers to go back and rethink how the different parts of the sentence are connected, which augments their cognitive processing load. A parallel sentence in ICEJAC-Purpose is good for reference, with a similar description of the research subject and objectives that maintains a succinct and lucid structure: "To assess the behavior of sand-silt mixtures, strain-controlled monotonic triaxial **tests were conducted** on sand-silt mixtures of specimen size 71 mm in diameter and 142 mm in height at various relative densities but same isotropic effective confining pressure of 100 kPa." It is a lengthy sentence as well. What distinguishes it is that the subject "tests" and the predicate "were conducted" are positioned adjacent to each other, notably enhancing the sentence's clarity and readability. Therefore, in the translated version of Case 2, a slight adjustment in word order should be made by placing "were carried out" ahead, following "studies", thus creating a more streamlined sentence structure and facilitating more efficient information delivery.

4.3 CCEJAC-Method

The Method move of academic papers in civil engineering often involves a precise and detailed description of experimental procedures. Its communicative purpose is to provide ample information about how the research was conducted, enabling readers to understand and evaluate the effectiveness and reliability of the study, and ensuring that the experiments can be

replicated. Due to the high frequency of verbs in the Method move, sentences often carry a substantial amount of information, leading to potential issues of excessively lengthy sentences. The following is an example of the description of experimental procedures in the CCEJAC-Method:

Case 3 分析了节理岩体的剪切力学特性劣化规律和劣化机理，并分段**建立了**考虑热湿循环作用损伤的节理砂岩剪切本构模型。

Translation Version: The degradation law and mechanism of shear mechanical properties of jointed rock masses **are analyzed**, and a shear constitutive model of jointed sandstone considering thermal wet cycling damage **is established** in sections.

CCEJAC involves a large number of passive voice constructions. As noted by Deng and Zhang (2023: 443), "In English, long sentences are typically compound sentences with multiple subject-verb structures; when the subject word count of a passive voice simple sentence exceeds half the average sentence length, the sentence tends to be unbalanced with a heavy top and light bottom." For instance, in Case 3, the original sentence is a non-subject sentence commonly seen in Chinese. In the translated version, it has been converted into passive voice, with the longer subject and its modifiers positioned at the beginning of the sentence and the verb placed at the end. Consequently, the sentence goes off focus. Moreover, in the Method move, the most important part lies in details concerning experimental objects and research procedures. Adhering to the end-focus principle of English sentences, it is more appropriate to place them at the end of the sentence. In ICEJAC-Method, a similar sentence also introduces research procedures, revealing a significant distinction: "The paper first **derives** the nonlinear equation of motion for the coupled system and then **compares** the analytical solution with finite-element mode." An active voice is used in this sentence, with a balanced structure: "The paper first derives… and then compares…". The sentence is concise, with the most important information placed at the end to clearly stress the sequence of research operations. Additionally, based on corpus-driven research, as Lu (2009) demonstrated, the structure "we + verb + object" is prevalent in expressing research content, results, and methods in English abstracts of scientific papers. There is no need to avoid using the first person to demonstrate the articles' objective and scientific nature. Employing the active voice promotes a more direct and clear expression, facilitating closer engagement with the reader. Therefore, it is advisable to use the subject with an active voice in the translation of research articles, and to render it as "this paper + verb + the object of the study" or "we + verb + the object of study".

4.4 CCEJAC-Product

In the abstract of academic papers in civil engineering, the Product move often presents new models, methodologies, and research findings; showcasing experimentally collected data and newly discovered insights. It also explains the successful validation of hypotheses through experiments. The communicative goal of this section is to explicitly convey to the readers the innovative aspects of the research and its contribution to the existing knowledge while demonstrating the effectiveness and feasibility of the research findings through data analysis. Here is a translation example from CCEJAC-Product that mentions research results:

Case 4 原文：声发射测试结果表明：玄武岩在水力耦合作用下的裂纹启裂为张拉破坏，**在裂纹稳定扩展阶段以张拉破坏为主，以剪切破坏为辅**，且这些破坏均主要发生在岩石中部。

Translation Version: The acoustic emission test results show that the crack initiation of basalt under hydraulic coupling is tensile failure, **which is mainly tensile failure and supplemented by shear failure in the stable crack propagation stage**, and these failures mainly occur in the middle of the rock.

In the original sentence, the phrase "以张拉破坏为主，以剪切破坏为辅" emphasizes the primary and secondary relationship between two types of failure, with "tensile failure" as the predominant mode, and "shear failure" as a relatively minor one. Concerning academic words, "supplemented" means "to add sth to sth in order to improve it or make it more complete", which deviates from the intended meaning. Furthermore, at the syntactic level, the relative pronoun "which" lacks clarity in its reference, leading to semantic inconsistency. The subordinate clause introduced by "which" and the main clause should be parallel rather than subordinate, describing the two failure stages separately. It is recommended that it be translated as "During the stable crack propagation stage, tensile failure predominates,

with shear failure occurring as secondary", to enhance the clarity of its logic and structure.

4.5 CCEJAC-Conclusion

In the Conclusion move of the research articles' abstract in civil engineering, authors may present inferences drawn from the research results, provide a thorough explanation of the findings, and highlight the practical significance of the research outcomes, shedding light on their implications for future research in the field. The communicative purpose of this move is to reinforce the readers' comprehension of the research results and to provide valuable insights for further exploration. The following is an example in CCEJAC-Conclusion, where inferences are drawn from the research results:

Case 5 调控水泥水化产物生成规整有序的水化晶体形状，改善界面过渡区的形貌，填充内部裂纹的空间，修复孔隙的形貌特征是掺入 GO 影响珊瑚砂水泥结石体抗氯离子渗透性的主要原因。

Translation Version: CCEJAC-Conclusion: Regulating cement hydration products to form a regular and orderly hydrated crystal shape, improving the morphology of the interface transition zone, filling the space of internal cracks, and repairing the morphological characteristics of the pores are the main reasons that allow the incorporation of GO to affect the resistance of coral sand cement stones to chloride ion permeability.

In Halliday's systemic functional linguistics theory, the theme in a sentence, which typically appears at the beginning, conveys information about the topic, while the rheme delivers detailed information, serving as the sentence's focus. Case 5's translation version presents a complex theme spanning 35 words, composed of four verb phrases in the subject. Halliday (1967) suggests that new information, usually the informational focus, should be placed at the end of the sentence. However, in this instance, the theme contains new information, representing the core content of the sentence, while the rheme conveys old information, which does not comply with Halliday's (1967) perspective. Therefore, adjusting the word order of the translation can better stress the new information: "There are four main reasons that allow the incorporation of GO to affect the resistance of coral sand cement stones to chloride ion permeability: regulating cement hydration products to form a regular and orderly hydrated crystal shape, improving the morphology of the interface transition zone, filling the space of internal cracks, and repairing the morphological characteristics of the pores."

5. Conclusions

This study constructed two monolingual corpora: Chinese Civil Engineering Journal Abstract Corpus (CCEJAC) and International Civil Engineering Journal Abstract Corpus (ICEJAC). Each corpus was further divided into sub-corpora based on their respective move. A comparative analysis revealed the following: At the lexical level, CCEJAC tends to use more abstract academic words, while ICEJAC employs more specific ones and demonstrates a higher frequency and richness in the use of general academic bundles compared to CCEJAC. At the syntactic level, CCEJAC shows longer average sentence length and unbalanced sentence focus, while ICEJAC displays shorter average sentence length and more balanced sentence structures.

From the case analysis perspective in CCEJAC, faithfulness in conveying the original meaning does not entail rigid adherence to the literal meanings of the original vocabulary, avoidance of the addition of general academic bundles, and refraining from alteration of the sentence structures of the source language. The findings provide practical references for enhancing the quality of translated abstracts of civil engineering research articles.

References

Chen X. W., Wang L.L. (2018). Translation of Humanities and Social Sciences Essay Abstracts: Bottlenecks and Countermeasures. *Shanghai Journal of Translators*, (01), 22-27.

Coxhead A. (2000). A New Academic Word List. *TESOL Quarterly*, *34*(2), 213-238.

Deng L.J., Zhang Y.Y. (2023). Use of Passive Voice in English Abstracts of Chinese Scientific Journals: Investigation and Analysis. *Chinese Journal of Scientific and Technical Periodicals*, *34*(04), 439-445.

Gardner D & Davies M. (2014). A New Academic Vocabulary List. *Applied Linguistics*, *35*(3), 305-327.

He D, et al. (2008). A Statistical Comparison and Analysis of English Abstracts in Chinese and English Scientific and Technical Periodicals. *Chinese Journal of Scientific and Technical Periodicals*, *19*(03), 404-407.

Hu J.M., Leng B.B. (2023). The Classification and Analysis of Typical Translation Errors of EST Texts in MTPE Mode. *Chinese Science & Technology Translators Journal*, *36*(01), 23-26.

Huangfu W.H., et al. (2012). Genre Analysis of English Abstracts in Academic Papers on Civil Engineering Written by Chinese Scholars. *Journal of Southwest Minzu University (Humanities and Social Science)*, *33*(S1), 205-209.

Hyland, K. (2000). *Disciplinary Discourses: Social Interactions in Academic Writing*. London: Longman.

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*(1), 4-21.

Liu L.J. (2015). Explicit Translation of Medical Paper Abstracts Based on Background Knowledge. *Chinese Science & Technology Translators Journal*, *28*(01), 17-20.

Li L., Fan W.Q. (2014). The Aesthetic Perspective of Translating Medical Paper Abstracts. *Chinese Science & Technology Translators Journal*, *27*(02), 12-15.

Lu Y.W. (2009). Personal Pronouns and Voice in English Abstracts of Scientific and Technical Papers: A Corpus-Based Study. *Chinese Journal of Scientific and Technical Periodicals*, *20*(06), 1167-1170.

M.A.K. Halliday. (1967). Notes on Transitivity and Theme in English. *Journal of Linguistics*, 204.

Nation, P. (2013). *Learning Vocabulary in Another Language (2nd ed.)*. Cambridge: Cambridge University Press.

Song D.H. et al. (2020). Semantic Feature Dictionary Construction of Abstract in English Scientific Journals. *Library and Information Service*, *64*(06), 108-119.

Sun W.H. (2015). On the Writing Practice and Principles Applied to the English Abstracts of Civil Engineering Papers. *Overseas English*, (14), 134-136.

Wang X.Q., Liu J.D. (2024). A Study on Types and Development of English Word List: Key Issues and Future Trends. *Foreign Language World*, (02), 90-96.

Xia L. et al. (2022). Machine Translation and Post Editing of English Abstracts in the Context of Artificial Intelligence. *Acta Editologica*, *34*(04), 396-401+406.

Xu W.J., Chen Z. (2019). English Abstracts of Civil Engineering Papers Based on IMRD Analysis. *Journal of Heilongjiang University of Technology (Comprehensive Edition)*, *19*(01), 100-105.

Yi L.M. (2008). A Genre Analysis on Abstracts of Civil Engineering Theses. *Journal of Chongqing University of Science and Technology (Social Sciences Edition)*, (03), 164-165.

Zhang J.H. (2011). Text Function and Translation of P. E Paper Abstracts. *Journal of Anhui Normal University (Natural Sciences)*, *34*(06), 602-606.

Zhang M, et al. (2024). JADE-DB: A Universal Testing Benchmark for Large Language Model Safety Based on Targeted Mutation. *Journal of Computer Research and Development*, *61*(05), 1113-1127.