

Research on the Efficiency of AI-assisted Translation in the Overall Lexical Level of the Target Language from the Perspective of Contrastive Rhetoric

Kaicheng Zheng, Guifang Wang
Zhejiang Yuexiu University, Shaoxing, China

How to cite this paper: Zheng, K., & Wang, G. (2025). Research on the Efficiency of AI-assisted Translation in the Overall Lexical Level of the Target Language from the Perspective of Contrastive Rhetoric. *Social Sciences and Humanities*, 2(1), 35-43. ISSN Print: 3104-4328; ISSN Online: 3104-4336. <https://doi.org/10.63313/SSH.9033>
Published: 2025-10-16

Copyright © 2025 by author(s) and Erytis Publishing Limited.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Abstract

Guided by the theory of English rhetoric, this study first identifies its research objects: the official English translation of the Constitution of the People's Republic of China from the National People's Congress website (referred to as "the original constitutional translation") and the AI-aided English translation by iFlytek (referred to as "iFlytek AI-assisted translation"). Then, based on the three principles of word choice in English rhetoric theory —Purity, Propriety, and Precision, it discusses the diction characteristics of AI-assisted translation through analyzing lexical data from the two-translation corpus, focusing on the dimension of overall lexical differences. Ultimately, through comparative analysis, the study reveals the effectiveness and limitations of AI translation in legal text processing while clarifying its translation effectiveness in the field of legal English.

Keywords

English Rhetoric Theory; AI-assisted Translation; Three Principles of Word Choice; Overall Lexical Differences

1. Introduction

In recent years, the era of artificial intelligence (abbreviated as "AI") is continuously developing and gradually transforming to thrive. Meanwhile, the traditional translation industry is undergoing a new round of innovation, which is driving its prosperity to some extent. For example, some scholars have proposed that "ChatGPT has pressed the "Accelerator Button" for the transformation of translation technology and foreign language education" (Wang Lifei, Li Zhao, 2023), and some also have pointed out that with the passage of time, the collision and integration between AI-assisted translation and human translation are becoming increasingly intense. Therefore, on the one hand, as a specialized communicative behavior that combines

professional knowledge in the legal field and a cross-language system, legal translation has been recognized as an important branch with a high professional threshold and an indispensable part of the entire translation system. It is no doubt that the strict requirements of legal translation for the preciseness, accuracy, and appropriateness of language rank at the top among all types of translation. On another hand, driven by the evolution of the times, AI-assisted translation is also rapidly expanding into the field of legal translation, meanwhile, with the increasingly frequent interaction between AI-assisted translation technology and the field of legal translation, the issues such as whether AI-assisted translation will have a substantial impact on the translation field and whether it can match the standards of human translation from the perspective of legal English specialization in the intelligent translation of laws, regulations, and relevant legal documents have gradually drawn the in-depth consideration of many scholars. Therefore, in order to further explore the actual effectiveness of AI-assisted translation in the field of legal translation, from the rhetorical perspective of legal translation, this paper takes the fundamental law of the People's Republic of China—the Constitution of the People's Republic of China, as the research object for translation. Through analyzing and comparing the differences of Core High-frequency Words between the official English translation of the Constitution and the version of iFlytek translation, this study explores the actual effectiveness of AI-assisted translation in the field of legal translation and reveals the potential impact trend of the development of AI technology on legal translation.

2. Overview of the Origin of English Rhetoric

The most important rhetorical concepts in West countries mainly originate from Plato's (427-347 BC) Philosophical Approach to Rhetoric "ethical/moral rhetoric", which was put forward in contrast to Socrates' (469-399 BC) dialectic (dialectic/dialektike) (MacDonald, 2017: 4) (Schiappa, 2017: 34). Compared with the scope of rhetorical research in later generations, its earliest form originated from the "practical rhetoric" in Aristotle's (384-322 BC) Rhetorique. His functional definition of rhetoric is "persuasion"—aiming to use means such as "logos" [logos, logical proofs or reason, corresponding to "reasoned discourse" (rational discourse), Schiappa, 2017: 38], "ethos" [ethos, ethical proofs or credibility, corresponding to "scientific persona" (scientific role), MacDonald, 2017: 22], and "pathos" [pathos, pathetic or emotion, corresponding to "artistic representation evoking pity or compassion" (representation that evokes pity or compassion), Merriam-Webster Dictionary], to logically, theoretically, and normatively interpret the content of the object, making it more authoritative and persuasive in the public domain. Aristotle once said, "If language cannot convey meaning correctly and clearly, it fails to fulfill its function" (Corbett, 1971: 422).

In the field of legal translation, whether the choice of words is appropriate and precise will all greatly affect the professionalism and accuracy of the expression of the

target language in the translation processes and results. It is also an important indicator for detecting whether AI-assisted translation has practical efficiencies in the field of legal texts translation. Therefore, how to ensure the accuracy and appropriateness of word selection in legal texts translation? After long-term research on various language texts and semantic communication methods, rhetoricians have summarized three principles or standards—the three principles of English diction, namely Purity, Propriety and Precision. These are also the crystallization of the wisdom of English rhetoric applied in the field of translation.

3. Connotations and Explications of the Three Principles of Word Choice

This theory is derived from the book *Lectures on Rhetoric and Belles Lettres* by the famous rhetorician Hugh Blair in the 18th century. The book states that words and sentences to make the translated text clear must possess three characteristics: Purity, Propriety and Precision (Blair, 1809: 167). John Sabine (1807: 44) also wrote in the book *A Guide to Public Speaking* that “in word-selection, we must abide by the principles of Purity, Propriety, and Precision.” Edward P. J. Corbett (1971: 423), a well-known contemporary rhetorician, clearly stated in his book *Classical Rhetoric for the Modern Student* that when choosing words for expression, one must always keep in mind the Three Principles of Word Choice in rhetorical theory.

In legal texts translation, since the target language was widely emphasized that must accurately and precisely express the content and the meaning of the original text, whether in single words or phrases, it cannot work without the Three Principles of Word Choice in English rhetorical theory. However, in the wake of that, there is another issue: how to implement the “Purity, Propriety and Precision” of the three principles of rhetorical diction in legal texts translation? Blair provided a concise summary of the concept of purity. According to his viewpoint, Purity refers to “employing words or structures that are in line with our linguistic conventions. It involves refraining from using words or phrases originating from other languages, as well as avoiding words or phrases that are archaic or inconsistent with proper linguistic authority.” In brief, in the process of legal translation, the words chosen for the target language should have the authority of expression, the rigor of professional terms, and the expressive power of appropriate wording.

Referring to “the selected words should conform to the established usage and can appropriately convey the meaning we intend to express” (Blair, 1809: 167). Propriety is also clearly mentioned by Edward that word-selection should conform to “Theme, Purpose, Occasion, and Audience” (Edward P. J. Corbett, 1971: 426). That is, in the semantic conveyance and expression of the original legal text, the target language should be used in a standardized manner to appropriately express the original meaning and theme of the text. Meanwhile, the adaptability of the target language to the intended audience and occasion should also be considered meticulously.

Finally, according to Blair's (1809: 169) elaboration, when it comes to the the main connotations of precision here are three key points as follow: (1)The words chosen in the target language fully express the author's original intention and viewpoint; (2)The words used comprehensively express the author's meaning and thoughts; (3)While expressing the author's meaning, the words used also convey the author's implied meaning, or even more than the author's implied meaning. In brief, in legal texts translation, the principle of precision emphasizes that each term in the legal text has a specific legal concept.in the process of translating legal text, it is necessary to strictly correspond to the professional vocabulary of the target legal system, eliminate the interference of colloquial translations in everyday language and avoid unclear expression caused by semantic ambiguity.

Then, from which dimensions of the translated text should we explore and analyze the actual effectiveness of AI-assisted translation from the perspective of rhetoric? As it known to all, words are the basic elements that make up sentences and texts and their importance is as inseparable as the relationship between roots and trees, water and fish while the relationship between legal translation and linguistics is as closely intertwined as the relationship between the day-night alternation and the movement of the sun and the moon. Therefore, when comparing different versions of legal translations, the core element remains the choice of vocabulary in the target language. And in legal texts translation, the examination of the words used in the target language is often the primary criterion for the efficiency evaluation. Thus, the analysis of the two versions of the translated text and the actual effectiveness of AI-assisted translation will be studied and discussed from the characteristics of the overall lexical differences in the translated text.

4. Research Design

The traditional method of word-frequency comparative analysis in translated texts is often constrained by the large base number and complex variety of text vocabulary and that' s the reason why It is difficult to efficiently conduct a comprehensive and centralized analysis of different translated versions while the results often fail to achieve a visual and intuitive effect as well. Therefore, in the research process of this paper, the AntConc Monolingual Corpus Retrieval Software is adopted for its precise and efficient vocabulary retrieval functions and powerful corpus-building methods. With the efficiency of its comprehensive coverage of text content and highly visual lexical data analysis function, a monolingual corpus of the translation version of the Constitution from the website of the National People's Congress of China (Corpus of Constitution of China National Congress website, hereinafter abbreviated as as CNC corpus) and a monolingual corpus of the translation version of the Constitution by iFlytek (Corpus of Constitution of IFlytek Translation, hereinafter abbreviated as IT corpus) are established. After comparing the high-frequency vocabulary data, through the statistical chart of high-frequency word data in the corpus , the top

twenty words in frequency in the CNC corpus and the IT corpus can be seen. Four indicators, namely Word Frequency (Freq_Tar, used to directly reflect the intensity of the target word's use in the translated text), Text Distribution Range (Range_Tar, used to reflect the distribution range of the target word in the translated text), Likelihood (Likelihood Keyness, used to test and calculate the significance efficiency of the target word in the text), and Effect (Effect Keyness, used to reflect the importance of the target word in the translated text) are leveraged to analyze the two translated versions.

5. Analysis of Core High-frequency Words on the Characteristics of Overall Lexical Differences

The scope of data results is selected within high-frequency words that show a high degree of overlap in core concepts in the two versions of the translated text. For example, "people", "state", "congress", "national", "article", "republic", "committee" and other commonly used proper nouns in legal translation are all ranked among the top ten in the high-frequency word list. Next, guided by the Three Principles of Word Choice in English rhetoric, a specific analysis will be conducted on the overall vocabulary difference characteristics of the translation from the four indicators in the visual data table.

5.1. Word Frequency

By comparing the word frequency data of the IT corpus and the CNC corpus, it is found that there is no significant difference in the word frequency data of the key words of the core concepts of the Constitution in the two translated texts. This phenomenon can also reflect that in the process of translating the Constitution text, iFlytek Translation adheres to the standard system of Constitution text translation in the use of core vocabulary. There is no interference from foreign languages or non-standard translation methods in the main core content. However, from the individual lexical differences, it can be found that there are three groups of words with slightly higher frequencies in the IT corpus, such as "people" (IT = 377, CNC = 376), "China" (IT = 97, CNC = 97), and "constitution" (IT = 26, CNC = 25). Compared with the standard translation on the website of the National People's Congress, the slightly higher use frequency of individual words in the iFlytek translation indicates that in the translation process, it attaches importance to the authority of professional vocabulary used in the transformation of the original language and the target language. The slight differences in word frequency do not lead to the introduction of impure words or deviation of core vocabulary. From the perspective of purity analysis, this meets the requirement of the equivalent conversion efficiency between the original language and the target language commonly accepted in legal English. Among the top twenty high-frequency core words in the IT corpus, only three core words (people, China, institution) have slightly higher word frequencies than the corresponding

words in the CNC corpus. This situation indicates that there are some differences in the use intensity of some core vocabulary in the complete translation compared with the CNC corpus, but the differences are within a reasonable range, that is, the word frequency meets the purity standard of Constitution translation. Therefore, in the process of translating the Constitution text, iFlytek Translation does not deviate from the original content and core meaning of the Constitution, which conforms to the principle that in the process of legal texts translation, the original text and the translated text should follow the overall consistency in terms of words and expressions in content and meaning.

5.2. Text Distribution Range

In addition to the word frequency data of core vocabulary, the equally important one—vocabulary text distribution range data, is often used to measure the coverage of the target word in the translated text. By comparing the IT corpus and the CNC corpus, it is found that the distribution ranges of the top 20 Core High-frequency Words are all 1, that is, the coverage of the core terms corresponding to these words in the two translated versions has reached a high degree of consistency. This indicates that in the corresponding text partitions of the translated text, it appropriately and properly maintains the term logic and concept expression integrity of the core content of the Constitution. From the perspective of the principles of the three principles of rhetoric, the high consistency in the coverage of key core words in the iFlytek translation and the official standard translation on the website of the National People's Congress indicates that AI translation has met the accuracy requirements for the conversion of legal translation proper vocabulary in the target language. For example, the distribution range of “congress” (代表大会) in both versions is 1, indicating that this word is uniformly used in all Constitution articles including “the National People's Congress” and “local people's congresses at all levels”, avoiding the ambiguity and variability of the target language caused by differences in expression forms and language logical structures in different Constitution regulations. Conversely, if AI translation only mechanically analyzes the meaning of the word “congress” in legal provisions, it may choose synonyms “assembly” and “council” similar in meaning to it as substitutes in the conversion of target language words to meet the “Universal of Avoiding Repetition” principle in translation (Nitsa Ben-Ari, 1998: 68). However, from the perspective of the legal translation rhetoric field, the synonymous substitution of proper words violates the requirement of the fixity of target vocabulary in legal texts translation and the clarity of legal concepts and the rigor of legal terms that the corresponding vocabulary should convey. Therefore, from the analysis of the distribution range data of the Core High-frequency Words in the Constitution translation, it can be seen that iFlytek Translation has met the requirements of non-ambiguous reference of translation vocabulary and clear and non-ambiguous selection of special words in the principle of accuracy.

5.3. Likelihood

In legal translation, in addition to the clarity of word use and the integrity of target language expression, we also need to consider the highly rigorous requirements for the conversion of Core High-frequency Words in the target language in legal texts. Both in professional vocabulary and basic vocabulary, the equivalent target language words and expressions formed with the passage of time have a certain irreplaceability. The comparison of this indicator in the vocabulary analysis of the CNC corpus and the IT corpus is evaluated through the efficiency of Likelihood. The data of “Likelihood” in the corpus reflects the significance of the target vocabulary in a specific text quantified by a statistical model. From the perspective of “propriety” in the three principles of rhetorical diction, the “appropriate pairing” of words in Constitution text translation requires that the significance ranking of terms must be faithful to the conceptual hierarchy of the original text. By comparing the data of this indicator in the two corpora, it can be concluded that the keyness of “people” (IT=1451.498, CNC=1530.823) and “state” (IT=652.000, CNC=693.372) ranks among the top two. This indicates that the highest legal status of “people” and “state” in the iFlytek translation is fully interpreted, and the word use and expression frequency are equivalent to those of the standard translation. From the perspective of core concept expression, the high consistency of the Likelihood of the Core High-frequency Words in the two versions in the ranking (“people” > “state” > “national”) shows that iFlytek Translation can highly identify the core concept chain of “people-state-nation”. At the same time, it also reflects that iFlytek Translation can achieve unbiased matching and transformation of core concept words in the translation of the Constitution and avoid the subversion of constitutional logic caused by the inversion of Likelihood (such as the significance of “state” being higher than that of “people”). Combining the above analysis, it can be concluded that the word use of the iFlytek translation, in terms of the cognitive fit with the importance of the original text concepts and the accuracy of the selection of core term words, shows that the iFlytek translation version also meets the requirements of using the target language in a standardized manner in the principle of “propriety”, and accurately and appropriately expresses the original intention and theme of the constitutional legal text and takes into account the rationality of “the subject and occasion to which the target language of legal texts translation is oriented”.

5.4. Effect

Different from the data of likelihood used to measure the irreplaceability of Core High-frequency Words in the translated text, the data of Effect focuses more on measuring the importance of the target word in the translated text and its actual contribution to the expression of the theme and concept of the translated text. In legal texts translation, to ensure that the translated text can balance the two core indicators of “faithfulness to the original text” and “target language standardization”, the

target language words corresponding to Core High-frequency Words often go through the screening and testing of the “semantic logic faithful transmission” criterion. Therefore, the final translated words play an important role in identifying the core logic and main meaning of the original legal text. In the data of the two corpora, the data of “Effect” can directly reflect the actual semantic influence of the target word in the translated text. By comparing the overall efficiencies in the CNC corpus and the IT corpus, it is found that the differences of the top 20 Core High-frequency Words in the two translated versions are all less than 0.0005, and the overall semantic transformation efficiency shows a consistent trend. This feature reflects that in the process of translating the Constitution text, iFlytek Translation ensures that the overall word use of the target language can be strictly compared with the word-selection standards of text professional terms, meeting the requirement of each term has a specific legal concept in the principle of “accuracy”. At the same time, from the perspective of the adaptation of text content and actual application scenarios, the Effect of “vice” (IT=0.004/CNC=0.005) in the two translated versions is lower than that of “president” (IT=0.060/CNC=0.062). The numerical differences of these two Core High-frequency Words in the two corpora are highly similar, indicating that iFlytek Translation reflects the subordination requirement of deputy positions to principal positions, can meet the cognitive matching standard between the subject and the audience in Constitution text translation, is highly consistent with the public’s understanding of the hierarchy of national institutions, such as “the vice-president assists the president in governing the country”, avoids the semantic distortion of professional term translation words, ensures the accurate combination of legal relations with the actual political situation understanding, and reflects that iFlytek Translation can follow the semantic transformation requirement of “the adaptability of the target language to the subject and occasion” in the principle of “propriety” .

6. Conclusion

This paper mainly explores the actual efficiency of AI-assisted translation in legal texts translation from the perspective of rhetoric. Taking the AI-intelligent translation version and the standard version of the Constitution of the People’s Republic of China as research objects, through analyzing the relevant data of the top 20 Core High-frequency Words, it analyzes the characteristics of AI-assisted translation at the overall lexical level, and concludes that in the processing of legal texts translation, AI translation has demonstrated the efficient transmission of “Standardized, Normalized” and precise terms. In terms of the “Purity, Propriety, and Precision” of overall word use, it is highly close to the human standard. Its core efficiency is data-driven standardized reproduction, and this translation mechanism has significantly promoted the feasible development potential of AI translation in the professional field of legal translation. However, AI-assisted translation technology mainly relies on the combi-

nation of statistical models and big data analysis models, and its practical application in the field of legal translation still faces the challenges of insufficient innovation and professionalism, which means its future development still needs to promote its own evolution from mechanical reproduction to standardized innovation through technological changes such as the iteration of exclusive domain corpora, the embedding of rhetorical rules, and the integration and incorporation of multi-modal data, and ultimately achieve the “Organic Synergy of Technical Efficiency and Humanistic Rhetoric”, so that the field of legal translation can truly move towards the path of intelligent and efficient development.

Acknowledgements

This research was funded by Major Humanities and Social Sciences Research Projects in Zhejiang higher education institutions, Grant Number: 2024QN165. And it's also supported by Special Research Project of the Foreign Languages Association of Zhejiang Province (No. ZWZD2025022).

References

- [1] Wang, L. F., & Li, Z. (2023). ChatGPT presses the "accelerator button" for the transformation of translation and foreign language education. *China Tech*.
- [2] Wang, H. S., & Liu, S. J. (2024). Research on smart translation education: Concepts, paths and trends. *Shanghai Journal of Translators*, 148(3), 1-8.
- [3] Blair, H. (1809). *Lectures on rhetoric and belle lettres* (11th ed.). Strahamand Prston.
- [4] Corbett, E. P. J. (1999). *Classical rhetoric for the modern students*. Oxford University Press.
- [5] Liu, L. H., & Zhan, M. Q. (2007). On the word-choosing characteristics of legal English from the perspective of rhetoric. *Science & Technology Information (Academic Research)*, (18).
- [6] Xia, F. F. (2013). On the word-choosing of the English version of the Contract Law of the People's Republic of China from the perspective of rhetoric. *Science & Technology Information*, (18).
- [7] Hale, S. (2007). *Legal translation*. Routledge.
- [8] Yuan, F. (2023). Research on the acquisition of legal English vocabulary based on a self-built corpus of textbooks. *Modern English*, (20).