

Language-attached Transfer Restrictions in Neural Machine Translation: An Analysis of *Tao Te Ching*

Qing Wang

School of Languages and Literature

University of South China, Hengyang, China

Received: June 11, 2024

Accepted: July 12, 2024

Published: July 30, 2024

doi:10.5296/ijl.v16i4.22002

URL: <https://doi.org/10.5296/ijl.v16i4.22002>

This thesis is funded by the Scientific Research Project from Hunan Provincial Department of Education, titled “Research on the Context Reconstruction Mechanism for the English Translation of Chinese Traditional Philosophical Classics in Natural Language Understanding (22C0230)”, the Scientific Research Project from Hunan Provincial Social Science Review Office, titled “Enhancing Quality Evaluation of Human-Machine Translation through Natural Language Processing-A Case Study on Texts for Special Purposes (XSP2023WXC002), the Teaching Reform Program for Hunan Higher Education, titled “Research on the Teaching Model of Translation Ability Training for Non-English Majors from the Perspective of Digital Humanities (HNJG-2021-0642)”, and the “14th Five-Year” Educational Science Planning Projects of Hunan Province, titled “Research on the Innovative Cultivation Path of Translation Learners from the Perspective of International Communication Capacity (ND232162)”.

Abstract

This study focuses on the language-attached transfer restrictions in Chinese to English neural machine translation (NMT) of *Tao Te Ching*, an ancient Chinese philosophical text, which aims to identify and analyze the specific challenges and limitations encountered by NMT systems when translating linguistic features that are unique to the source language and how these challenges impact the accuracy and quality of the translated text in English. It contributes to the understanding of the complexities involved in machine translation of ancient Chinese texts and proposes strategies to mitigate the identified transfer restrictions, ultimately aiming to bridge the gap between machine translation capabilities and the richness of Chinese philosophical literature.

Keywords: Neural machine translation (NMT), Language-attached transfer restrictions, *Tao Te Ching*, Linguistic features, Cultural sensitivity

1. Introduction

Most translation studies are about language, but it is not about one language, but two languages, involving the comparison and transformation of different languages and cultures (Tan, 2017, p. 13). Machine translation is a process of transforming one natural language into another by using computers. Therefore, the problem of machine translation is also related to natural language understanding and processing, including word segmentation, annotation, word order, ambiguity, context and logical relationship. Whether these key problems can be effectively solved is directly related to the quality of machine translation system. Solving and dealing with language problems in machine translation is the primary task to achieve good machine translation (Wang & Ma, 2020, p. 145). From this point of view, the quality of machine translation depends on whether the lexical, syntactic, grammatical, semantic and contextual factors of the text can be identified, understood and processed by computers.

2. Language-attached Transfer Restrictions

Precedent studies have focused on the language-leveled restrictions in the process of machine translation, mostly about the ones related with language rules but seldom concern about the pragmatic and cognitive sides. While the problem of machine translation stems from the basic problems of linguistics, involving the conversion of linguistic knowledge such as grammar and semantics and trans-linguistic knowledge such as context, emotion and communicative intention, the fact is that the languages of Chinese and English originate from different linguistic systems, thus there should be some restrictions within or beyond language when transfer. **Then what are these specific transfer restrictions?**

2.1 Grammatical Features

How to maintain the original grammatical features such as articles, pronouns, positions, plurals, tenses and logical relations, etc. Different from manual translation, machine translation is more difficult because machine translation systems cannot think autonomously like human beings. In order to make the source text fully reflected in the translation, the premise is that machine translation systems can properly convert and adjust according to the grammatical rules of the two different languages. However, in terms of retaining the grammatical features of the original sentence, machine translation systems can only function according to the artificially predefined grammatical rules and features. Setting huge grammar rules and complex language features is just a task and goal that is difficult to accomplish and achieve by manpower alone. As far as Chinese and English are concerned, they belong to the Sino-Tibetan language family and the Indo-European language family respectively. Their language structures are different, and each has its own language use habits and characteristics. From the expression of the two languages, Chinese is vocabulary-attached language and English is letter-attached language. Lexical attachment language refers to that in the language system, the word form of vocabulary does not represent any grammatical meaning, and various language relations in sentences need to be reflected by adding special reference words. For example, the semantic relationship of quantity in Chinese must be expressed by adding quantifiers. Letter-attached language refers to the language system in which the word form of a word represents a certain grammatical form. Various linguistic relations in a

sentence can be reflected by changing the word form of a word, such as English.

The changes of gender, number, case and position in English can be expressed by suffixes and prefixes. In terms of grammatical features, the differences between Chinese and English are mainly manifested in Chinese functional words, English articles, word order and sentence patterns. Among them, functional words play a very important role in Chinese sentences. Without it, the complete meaning of sentences cannot be expressed. The difficulty of machine translation lies in the fact that some functional words such as auxiliary words and modal words in Chinese have no corresponding vocabulary in English language, such as “ma”, “le”, “zhe”, “guo”, “di” and “ah”. The meaning of such functional words in sentences is mainly expressed through the change of English grammar. There are several situations listed as follows, such as using auxiliary verbs, be verbs, modal verbs, etc. to express general questions, like “Are you a student?”, “Can you speak English?”, “Do you know him?”, etc., and utilize the change of tense to reflect the time of the event, such as “He is reading.”, “I have finished my homework.”, “She will go abroad.” Also English language generally begins with “what” and “how” to express Chinese exclamatory sentences with the “how” tone, such as “What a beautiful girl.”, “How beautiful the girl is.” In addition, the handling of sentence patterns is an indispensable approach in the process of machine translation conversion. For example, English has unique sentence structures for inverted sentences and passive sentences, which must be specially handled. “One of the basic characteristics of modern Chinese grammar is that word order plays a very important role (Zhang, 2016, p. 92)”. Due to the lack of rich morphological forms, various grammatical forms, grammatical relationships, and linguistic meanings in Chinese can only rely on changes in word order, and different word orders will form a variety of corresponding sentence patterns. In view of the different language expressions of Chinese and English, it is a necessary prerequisite to input the grammatical rules of Chinese and English languages into the Chinese-English machine translation system, and it is also the primary condition to make the translation as faithful to the original text as possible.

2.2 Sentence Segmentation

How to segment words and phrases within sentences in a text language. Segmentation of words and sentences is an initial task in natural language processing, playing a decisive role in whether computer programs can correctly recognize and understand the meaning of sentences. Saussure (1980) believed that “an element derives its value from being related to another element, either preceding or following it, or both”, indicating that the grammatical forms produced by word combinations and collocations reflect various linguistic relationships. Word collocation is a linguistic phenomenon where words co-occur (Halliday, 1976, pp. 284-292), generally influenced by lexical, grammatical, semantic, logical, and habitual collocation rules. Based on whether they follow universal collocation rules, they can be classified into fixed collocations, conventional collocations, and deviant collocations (Zhou, 2008, p. 2). Deviant collocations, in contrast to fixed and conventional ones, refer to lexical combination relationships that do not follow conventional lexical, grammatical, semantic, logical, and habitual collocation rules. Unlike normal collocations that reveal the habitual co-occurrence patterns between words, deviant collocations uncover both explicit and

implicit patterns among words. Since deviant collocations are neither constrained by co-occurrence probability nor related to general collocation rules, and are influenced by non-linguistic factors such as social, ethnic, cultural, aesthetic, and psychological elements, whether mastering the rules of deviant collocations or not presents a challenge for natural language processing. It also brings uncertainties in automatic segmentation for computer systems. Moreover, for natural language processing, the segmentation of language fragments is crucial to understanding the entire sentence's meaning; different positions of delimiters convey different meanings. For instance, in “*拿上衣架*”, “*衣*” can form “*上衣*” with the preceding context or “*衣架*” with the following context. However, based on the semantic relationship of the context, the division here should be “*拿上/衣架*”. Furthermore, the function of delimiters can more clearly reflect the specific content described by the sentence, as in this example, the object of the action is *衣架*, but not “*衣服*” or “*架子*”. Similarly, in this sentence “*我/明天/要去/巴黎*”, the use of delimiters clarifies the subject, time, and location of the event, which is very significant for machine translation. It not only helps computers better understand the semantic relationships in the original text but also enables more accurate translation into the target language. However, the horizontal combination relationships between linguistic units cannot solely rely on habitual usage or literal meanings. The hidden deep meanings and related cultural background knowledge should also be considered. Therefore, resolving the issue of language fragments segmentation during machine translation is a necessary prerequisite for improving translation quality.

2.3 Lexical Ambiguity

How to eliminate lexical ambiguity and select the precise meaning. The most fundamental issue in natural language processing is the identification of multiple meanings for a given word (Zhang, 2004, p. 27). Lexical ambiguity refers to the phenomenon where a single word has multiple meanings, which is common in natural languages. The factors causing lexical ambiguity are complex because words prone to ambiguity inherently contain multiple grammatical meanings that do not fit all linguistic relationships, and the specific meaning of a word is determined by the context with which it collocates. In the process of machine translation, selecting an incorrect word sense can affect the fidelity of the translation, failing to accurately convey the original content, thereby rendering the machine translation ineffective. Therefore, word sense disambiguation is a hot topic in the field of machine translation and one of the challenges that must be addressed by machine translation systems. Whether in English or Chinese, polysemy is rampant, for example, the word “*watch*” can refer to “*wristwatch*”, “*lookout*”, “*observe*”, or “*guard*”, and the Chinese character “*盛*” can mean “*flourishing*”, “*grand*”, “*common*”, or “*profound*”. Lexical ambiguity is not confined to nouns and verbs alone, adjectives, pronouns, prepositions, and conjunctions can also cause ambiguous phenomena. The true meaning of these words depends on their linguistic context, and the correct linguistic meaning must be inferred from numerous possible meanings based on contextual information. Additionally, the inherent meaning of the word, common sense, and world knowledge are important reference conditions for eliminating lexical ambiguity. The difficulty of word sense disambiguation lies in the irregularity in the distribution of polysemous words, which presents a particularly challenging task for computer in natural

language processing (Feng, 2012, p. 96). Among the vast number of polysemous words, some can collocate with certain vocabulary to express the same idea, such as “the significance of success” and “the significance of failure”, where although the collocated words have opposite meanings, they express the same concept, namely the significance and impact brought about by events. Of course, there are also words that, when combined with other words, form a linguistic relationship but express multiple different meanings, such as “an aggressive man”, which can be understood as “a person with ambition” or “a person with a strong tendency to attack”. For these complex polysemous phenomena in language, computer programs cannot easily understand the specific meaning of the language and select the appropriate linguistic units as humans do. Consequently, the task of disambiguation before translation becomes particularly crucial for computers.

2.4 Pragmatic Analysis

How to conduct pragmatic analysis of textual language sentences. Syntactic analysis, semantic analysis, and pragmatic analysis are the basic procedures for computers to understand sentences and are also the foundational prerequisites for ensuring the correctness of machine translations (Luo, 2001, p. 267). Pragmatic analysis is a deeper level of linguistic analysis that follows syntactic and semantic analyses, revealing the relationship between discourse content and real-world entities and their attributes within a certain context. It manifests the literal and implied meanings produced by language use through a dynamic structure of signification. Unlike the meaning studied in formal semantics, pragmatics focuses on revealing the specific meanings reflected when language is used in certain contexts. Similarly, computational pragmatics studies the relationships between discourse and action, discourse and text, as well as discourse and the time, place, and surrounding environment of conversation from a computational perspective (Jurafsky, 2001, pp. 1-36). It can be said that the context of language use is a key means to understand specific information and communicative intentions of language, reflecting the true thoughts expressed by language to some extent in the entire translation process. The relationship between the meaning of language and its context is like “the pearl in the crown”, that is, the same pearl, placed in different locations, exhibits different values and meanings. For example, in Context A, “Unbelievable” is said after a person with a fear of heights completes a skydiving action, while in Context B, “Unbelievable” is said when a mother sees her child smearing paint on themselves. The same word expresses entirely unrelated meanings in these two different contexts. During pragmatic analysis of sentences, besides contexts, speech acts, conversational implicature, pragmatic presuppositions, deictic information, and discourse structure are also needed to interpret. Interpreting these textual contents often requires the use of human logical thinking abilities for judgment and reasoning, to uncover knowledge and characteristics beyond linguistic components and symbol units, so as to truly reflect the actual intention of the language. For instance, in the sentence “Lisa has a big mouth”, there could be two meanings, one refers to Lisa’s physical characteristic, which means she has a large mouth, expressing whether she looks nice or not, while the other implies that Lisa cannot keep secrets, talking about everything to everyone, expressing whether it seems good or bad. Thus, the difficulty in machine translation lies in the fact that computer systems do not possess

human-like logical reasoning abilities. For the implications beyond the literal meaning of sentences, computer systems may not always recognize them, and sequencing and coding are the basic approaches for natural language processing. The conversion between textual language sentences is achieved through rational mathematical models. For abstract concepts such as non-linguistic components, they must first be converted into computable data representations before they can be recognized and understood by computers. In the preliminary work of pragmatic analysis, defining abstract hypertext components, concepts, and categories solely relying on computers can be very challenging, because practical operations still depend on humans. This also accounts for the fact that the introduction of pragmatic knowledge into computer systems has developed slowly.

2.5 Stylistic Features

How to map the stylistic features of the source language. A complete sentence, besides having its own grammatical rules, also reflects the speaker's language style and speech characteristics to varying degrees. In Gutt's view (1991), these stylistic features are "communicative clues" that can lead readers to truly comprehend the original author's thoughts and intentions, making it extremely important and valuable to preserve these elements. Therefore, in the process of language transformation, some artistic styles and linguistic characteristics of the original text should also be maximally reflected in the target language. Among Fu Yan's three translation criteria of "faithfulness, expressiveness, and elegance", the modern meaning of "elegance" refers to the reproduction or re-creation of the stylistic and characteristic features of the original text's sentences in the translation. However, the quality of machine translation currently cannot fully achieve the level of "elegance", as machine translation systems have limited understanding of natural language and can only implement the mapping from source language text to target language text based on human-preset grammatical rules and semantic knowledge through computer programs. During the translation of ancient poems by computers, there is no difference in the approach compared to the translation of ordinary texts, it also involves finding matching translation units or fragments from dictionaries, corpora, semantic knowledge bases, and other language material storage libraries for both the source and target languages. However, revealing implied hypertext such as rhyme, rhythm, and rhetoric from the language knowledge base through machine translation is more challenging. For example, comparing the manual translation with the machine translation of the Chinese ancient poem "静夜思":

Original Text:

床前明月光，疑是地上霜。举头望明月，低头思故乡。

Xu's Translation:

A Still Night

Before my bed a pool of light,

O can it be frost on the ground ?

Looking up, I find the moon bright;

Bowing, in homesickness I'm drowned.

Baidu Translation:

In the Quiet Night,

The bright moon in front of the bed

is suspected to be frost on the ground.

Raising my head, I see the moon so bright;

Withdrawing my eyes, my nostalgia comes around.

Youdao Translation:

A tranquil night,

There was moonlight before the bed,

and it was frost on the ground.

Look up at the moon,

how to think of home.

Comparing the above translations, it reveals that the greatest difference between professional translators and machine translators lies in the former's ability to fully leverage human subjectivity and creativity, achieving a certain level of aesthetic and appreciative quality in the translation. In Yuanchong Xu's translation, there is an abab rhyme scheme that makes it pleasing to read aloud, and the sentences are poetic with novel sentence structures. In contrast, while machine translation can render the general content of the verses, its wording is overly direct, the sentence structures are conventional, and it lacks creativity. As is well known, the fundamental reason for the gap between manual and machine translation is that machine translation follows rules preset by humans, thus it is not difficult to conclude that computers operate in a state of thoughtlessness and lack of initiative during the translation process. The challenge for machine translation is that if every linguistic artistic feature of each sentence were to be programmed into the system, it would require considerable human and material resources. On the other hand, if only a macro-level control is exercised, some local linguistic rhetorical techniques cannot be presented through the computer. Nida's view on machine translation versus human translation is that "memory does not equal emotion, nor does speed replace beauty" (1964, p. 264). Machine translation systems lack human appreciation and creative abilities, relying solely on random program conversions for language artistic style features that cannot be matched does not allow the translation to truly reach a certain level of

appreciation. Therefore, before machine translation systems can automatically recognize the artistic aesthetic characteristics of language in texts, it will be quite challenging for machine translations to meet the standard of “elegance”.

2.6 Common sense and World Knowledge

How to recognize the common sense and world knowledge of language texts. To improve the translation quality of machine translation systems, it is not only necessary to address issues related to the language itself but also to understand some content and information hidden behind the language texts. Injecting a large amount of common sense such as world knowledge and encyclopedic knowledge into the machine translation system is an indispensable step for computer programs to understand and generate natural language. This is also due to the considerations of the different cultural backgrounds of source and target languages. Two different ethnic cultures inevitably have different understandings and perceptions of certain things or phenomena. For example, the implication of “dragon” symbolizes evil and ugliness in Western countries, playing the role of rebels and fools, while in China’s history and culture, it is a symbol of peace and harvest, a protector and embodiment of good fortune for the Chinese people, from which most Chinese parents’ wish for their children to become successful originates. Furthermore, there are significant differences in the way addresses are used between China and the West. Chinese addresses differentiate by gender, seniority, and status, such as mother, aunt, uncle’s wife, and maternal aunt, stemming from the feudal hierarchy that emphasized differences in rank and focused on blood relations. In contrast, English has a limited number of addressing names, often generalizing relationships of the same attribute into one term, such as Chinese cousins being translated as “cousin” in English. This is because English culture generally does not distinguish gender differences in titles and does not care about seniority in addressing elders, even allowing the use of their names directly. In the process of machine translation, lacking the recognition of social, historical, cultural knowledge, and common sense of language can result in translations that do not faithfully adhere to the original text. Computer systems do not possess the same cognitive concepts about things and phenomena as humans during the translation process. For humans, the cognitive process is a summary of experience and accumulation, whereas for machines, various rules, algorithms, and corpora are the cognitive foundation and reference during the translation process. Compared to the human brain, what computers lack is an inherent set of value systems that humans possess. Therefore, in the translation process, machine translation systems indirectly acquire conventional knowledge and general concepts hidden behind language texts from knowledge created by humans, such as “China” and “china” referring to the country and the artwork respectively. The difference between these two words lies only in the capitalization of the initial letter but expresses different meanings, yet they are related, that is, “China is a country rich in china”. Computers must recognize the difference and specific usage between these two from the massive corpus knowledge stored in the system to obtain their conventional meanings in language usage. This passive approach to machine translation means that the quality of machine translation excessively depends on the breadth and depth of internal language knowledge storage within the system.

3. Examples of NMT Use

Machine translation faces numerous linguistic challenges, especially when translating Chinese classics like *Tao Te Ching* into English. These challenges often relate to the issues of lexical equivalence, syntactic structure, cultural adaptation, rhetorical devices, stylistic complexity, etc.

3.1 Example 1

ST: “上善若水。水善利万物而不争，处众人之所恶，故几于道。”

NMT: “The highest good is like water. Water benefits all things and does not compete with them; it dwells in places that people despise, thus it is close to the Dao (Way).”

3.1.1 Lexical Equivalence

The term “上善” refers to the highest form of goodness, while “水” often symbolizes softness and compliance with nature in Chinese culture. The machine translation renders “上善” as “the highest good”, conveying the concept of superiority but possibly lacking the philosophical depth embedded in the original. Similarly, “水” is translated into “water”, preserving the original word, but its cultural connotations and symbolic meanings require additional explanation for full comprehension.

3.1.2 Syntactic Structure

Chinese syntactic structures differ significantly from English. Sentences in *Tao Te Ching* are often concise and rhythmic, and translating them into English necessitates adjusting the sentence structure to fit English grammar rules. In this example, the machine translation successfully transforms the concise style of the original into English while maintaining fluency and readability.

3.1.3 Cultural Adaptation

The term “道” is a central concept in Taoist philosophy, referring to the fundamental principle governing the operation of all things in the universe. The machine translation renders it as “the Dao (Way)”, adding “Way” in parentheses as an explanatory note. This is a strategy of cultural adaptation aimed at providing additional information for readers unfamiliar with the concept.

3.2 Example 2

ST: “道可道，非常道。名可名，非常名。无名天地之始；有名万物之母。”

NMT: “The Tao that can be told is not the eternal Tao. The name that can be named is not the eternal Name. The unnameable is the beginning of heaven and earth; The nameable is the mother of all things.”

3.2.1 Conceptual Implication

The machine translation maintains fidelity to the original text by keeping key concepts like “Tao” and “Name”. However, the nuanced philosophy of the “eternal Tao” and “eternal Name”

may be lost on readers unfamiliar with Taoist thought. The translation does not fully convey the profound philosophical implications of these terms.

3.2.2 Cultural Context

Phrases like “eternal Tao” and “mother of all things” are heavily influenced by Chinese cosmology and philosophy. Machines may struggle to accurately convey the cultural significance and nuance without additional explanatory notes.

3.2.3 Rhetorical Devices

The repetition and contrast as “可道”/ “常道”, “可名”/ “常名”, “无”/ “有” present in the original are effectively preserved in the translation, that is “The Tao that can be told”/ “the eternal Tao”, “The name that can be named”/ “the eternal Name”, “unnameable”/ “nameable”. This helps in retaining the rhythmic and philosophical depth of the original.

3.2.4 Stylistic Fidelity

The translation maintains a poetic style similar to the original, which is crucial for a work that straddles both prose and poetry. The translated sentences are simple and clear, yet they convey a sense of depth and mystery, aligning with the style of the source material.

3.3 Example 3

ST: “天长地久。天地所以能长且久者，以其不自生，故能长生。是以圣人后其身而身先，外其身而身存。”

NMT: “Heaven and Earth are enduring. The reason why Heaven and Earth can be enduring and long-lasting is that they do not live of themselves. Thus they can endure and last long. Therefore, the Sage places his person in the background, yet he is ahead of all others. He puts his person outside, yet it survives.”

3.3.1 Idiomatic Usage

The text contains idiomatic expressions such as “后其身而身先”, that is, placing oneself in the background yet being ahead, which cannot be translated word for word but require an understanding of the underlying concepts to convey the meaning accurately.

3.3.2 Philosophical Concepts

The passage includes philosophical ideas like “不自生”, that is not living for oneself, which need to be understood within the context of Taoist philosophy to translate them effectively. Machines may lack the cultural and philosophical nuance required to interpret these concepts correctly.

3.3.3 Conceptual Clarification

Terms like “圣人” have specific connotations in Chinese philosophy that may not directly map onto any single English word, thus requiring additional context or explanation for clarification.

3.4 Example 4

ST: “上德不德，是以有德；下德不失德，是以无德。”

NMT: “Highest virtue is not (the thought of) virtue, and thus it truly is virtue. Lowest virtue never loses (the thought of) virtue, and thus it is really no virtue at all.”

This machine translation attempts to address the differences in grammatical structures between Chinese and English by using brackets to indicate elements that are implicit in the original text but must be explicit in English.

3.4.1 Polysemy and Ambiguity

The term “德” has multiple meanings and can vary significantly depending on context. Machine translation might struggle with accurately capturing its specific interpretation in a given context, leading to potential inaccuracies or misunderstandings. In this context, “德” primarily refers to an intrinsic moral quality or excellence that is closely related to individual behavior and cultivation yet transcends ordinary moral norms or external moral actions.

“上德” signifies the highest form of moral quality, so natural and internalized that the person possessing it is not aware of its presence. This virtue is formless, and it is not the result of deliberate pursuit but rather a manifestation of going with the flow and harmonizing with the Tao. Thus, “不德” in this context means unpretentious, effortless, and natural. Someone with the “highest virtue” does not need to deliberately display their virtuousness, and their virtue naturally emanates from within. True virtue is intangible and resides in one’s essence without relying on external recognition or self-promotion.

“下德” represents a lower level of moral quality. The phrase “不失德” here means a conscious effort to maintain or showcase one’s virtue, involving a deliberate adherence to moral codes. Although this behavior is moral, it is considered “无德” because it contains elements of pretension and self-congratulation. This indicates that true virtue is not achieved through external coercion or self-promotion but through the natural expression of one’s inner qualities.

Overall, while the machine translation renders the original text into English with regard to syntax, rhetoric, and style, it may still leave room for improvement in fully conveying the deep philosophical implications and cultural context inherent in *Tao Te Ching*.

4. Discussions and Conclusions

In specific behavioral patterns, the relationship between translation and language is equality or beyond. Translation achieves equivalent transformation between two languages, yet it involves more than just the transfer of internal linguistic knowledge, including the transmission of external linguistic knowledge such as society, culture, cognition, emotion, etc., making translation greater than language itself. Therefore, the success or failure of translation is inseparably linked to national culture, reader cognition, and the application of grammar and language habits. Firstly, a nation’s language is closely related to its culture, with each influencing and complementing the other. The language used depends on the cultural

context in which it is situated. “In a sense, translation is about translating cultures”(Fang & Fan, 2015, p. 175). The absence of cultural differences in translation can lead to a lack of intended pragmatic effects in the translated text, hence the transplantation and integration of different cultures between two languages is an important factor that cannot be overlooked in the translation process. Secondly, the communicative process of language is actually a process where bidirectional cognitive contexts influence and constrain each other (Zhao, 2016, p. 71). Due to the differences in cognitive contexts among readers of various nations, that is, “the subset of assumptions about the world that the hearer has” (Sperber & Wilson, 1995, p. 15), target language readers lack cultural schemas related to ethnically characteristic languages. Without explanatory translation, certain ethnic languages cannot achieve the intended contextual effects during the cognitive reasoning process, thus leading to cognitive errors. Furthermore, different languages vary greatly in terms of grammar usage and language habits. Taking Chinese and English as examples, Chinese vocabulary does not change for gender, number, or tense, whereas English vocabulary follows a strict set of morphological rules. English is a “tense language”, primarily expressing tense through overt verb conjugations, while modern Chinese mainly uses temporal nouns, adverbs, and other lexical means to express temporal meanings. Although the Chinese word “了” is gradually becoming a tense marker, it is still considered a “semi-tense language” (Park & Yuan, 2019, p. 448). Additionally, Chinese focuses on sentence topic, often neglecting sentence form for the described objects, such as omitting subjects, removing words of little significance, and eliminating repetition, retaining only the relevant main content such as atmosphere, plot, and background environment. In contrast, sentences describing objects in English follow strict grammatical and language use rules, emphasizing the completeness of sentence structure and retaining the basic subject-verb-object structure. It is these grammatical differences, cultural differences, cognitive differences, and thinking differences between different languages that make it more challenging for computer programs to process and understand natural language, thereby resulting in translation that is inferior to that of professional translators.

Fundamentally, the ultimate and ideal state of machine translation is to possess the ability to analyze, understand, and apply language like human translators (Gui, 2018, p. 301). However, the complex relationship between language and thought determines that natural language processing is not a simple matter. “Understanding written language is an extremely complex intellectual skill” (Wang, 1996, p. 81). Humanity has yet to fully comprehend the process by which the human brain recognizes, understands, and generates language, much less rely on computer systems to perform high-quality translations. “The translation process occurs within the human brain’s consciousness and cannot be directly observed. Texts written or spoken are the material objects of specific translation analysis” (Luo, 2018, pp. 233-239). Whether a computer, as a method of machine translation without autonomous consciousness, can recognize the various differences between textual languages will directly affect the outcome of machine translation. Therefore, the difficulties in handling issues such as lexical ambiguity, word collocation, linguistic style characteristics, contextual context, and background knowledge become obstacles to improve machine translation performance and enhance the quality of machine translation.

Before the 1990s, machine translation was predominantly rule-based, while afterward, it has tended toward data-driven approaches. Although the output of statistical machine translation and neural network machine translation based on large-scale corpora significantly surpasses that of rule-based machine translation, their implementation still requires adherence to basic grammatical rules of language. Regardless of how machine translation evolves, its aim is always to enhance the capability of machine translation systems. Clearly, machine translation systems cannot completely replace humans in performing translation tasks, and their assistance lies in sharing some of the routine workload of human translators, meeting basic translation needs such as simple conversational exchanges, multilingual media information retrieval, and basic understanding of literary works. For higher translation demands, such as achieving publishable quality, neural network machine translation based on deep learning can barely meet the standard, but other machine translation methods like direct method, transfer method, interlingua method, example-based method, and statistical method still fall short. Although the study of machine translation has achieved new developments and breakthroughs after experiencing infancy and bottlenecks, the goal of fully automated high-quality machine translation still has a long way to go. From the perspective of the machine translation process, linguistic issues are the most fundamental and critical problems hindering the full conversion between two languages. Only by overcoming obstacles at the lexical, syntactic, semantic, and pragmatic levels can machine translation truly achieve fidelity and fluency. Conversely, the degree of adequacy and acceptability of machine translations also reflects, to some extent, the computer's ability and effectiveness in processing and understanding natural language.

References

- Fang, M. Z., & Fan, W. Q. (2015). *A Course in Scientific and Technological Translation*. Shanghai Foreign Language Education Press.
- Feng, Z. W. (2012). *A Concise Course in Natural Language Processing*. Shanghai Foreign Language Education Press.
- Gui, Q. Y. (2018). *New German Translation for Beginners*. Tongji University Press.
- Gutt, E. A. (1991). *Translation and Relevance: Cognition and Context*. Oxford: Basil Blackwell.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Jurafsky, D. (2002). *Pragmatics and Computational Linguistics*. Oxford: Blackwell.
- Luo, P. (2018). A Review of Issues in Russian and Soviet Translation Theory Texts. In Zhang, C. (Ed.), *Collection of Essays on Foreign Languages and Literatures* (Vol. 8, pp. 233-239). Sichuan University Press.
- Luo, X. M. (2001). *A Comparative Study of Discourse Analysis in English and Chinese*. Hunan People's Publishing House.
- Nida, E. A. (1964). *Toward a Science of Translating*. Leiden: E J Brill.

<https://doi.org/10.1163/9789004495746>

Park, M. A., & Yuan, Y. L. (2019). Is Chinese a “Tenseless Language”? *Contemporary Linguistics*, 438-450.

Saussure, F. De. (1980). *Course in General Linguistics*. Gao, M K. Beijing: The Commercial Press.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition* (2nd ed). Oxford: Blackwell.

Tan, Z. X. (2017). *Translation Studies: The Quest and Development as an Independent Discipline*. Fudan University Press.

Wang, K. Z. (1996). *Natural Language Understanding—Can Computers Think?*. Harbin Institute of Technology Press.

Wang, Q., & Ma, X. (2020). Research on Machine Translation Quality Assessment Methods from the Perspective of Problem Awareness. *Hunan Social Sciences*, 144-151.

Zhang, W. X. (2016). *Grammar and Grammar Teaching Research*. Beijing Language University Press.

Zhang, Z. (2004). Preliminary Discussion on Machine Translation. *Chinese Science and Technology Translators Journal*, 23-28.

Zhao, H. (2016). *A Cognitive Pragmatic Study of Verbal Irony in Novel Communication*. Shandong University Press.

Zhou, C. L. (2008). *Research on the Deviation and Collocation of Word Semantics and Grammar*. Yunnan People’s Publishing House.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>)