

Research Article

The Impact of Big Data Technology on Phrase and Syntactic Coherence in English Translation

Xiaoyu Li 

School of Foreign Language, Jilin University of Architecture and Technology, Changchun 130111, China

Correspondence should be addressed to Xiaoyu Li; lixiaoyu@jluat.edu.cn

Received 28 May 2022; Accepted 5 June 2022; Published 25 June 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Xiaoyu Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the era of big data, data are ubiquitous. The proliferation of data leads to a surge in the demand for communication, which in turn promotes a surge in the demand for language services. Big data technology is a comprehensive technology, and its important feature is a more active technological factor, where technological development and technological innovation dominate, which will have an incalculable impact on the development of the translation industry. In the language service industry, much information that used to be difficult to quantify will be transformed into data for storage and processing, and a large number of complex items to be translated will gradually surface. Therefore, it is the general trend of the translation industry to stimulate and utilize the unexplored values hidden inside the data and develop the blue ocean of the language service industry. Traditional translation researchers are confined to the study of language and text and are not fully aware of it. The huge role played by translation technology in today's business environment, and the traditional translation theory can hardly describe and explain the new modern translation technology phenomenon and translation technology activities. Whether we are ready for it or not, the rapid development of translation technology in the era of big data will lead to significant changes in translation inquiry and translation teaching on a global scale. The integration and development of translation with cloud computing, big data, the Internet, and artificial intelligence are driving a series of innovations in the traditional translation model. The development of translation technologies and tools is breaking new ground, rapidly expanding into all aspects of the translation industry and triggering disruptive changes in the language service industry.

1. Introduction

Technological innovations, especially the rapid development of cloud computing, big data and artificial intelligence technologies have made people's lives more convenient; in the translation industry, they have promoted the emergence of many new translation technology models [1]. Mechanical translation has officially become an integral and important part of the history of translation and one of the highly regarded cutting-edge technologies. Automated translation technology is a technique for quickly converting written or spoken language into many different languages [2]. For a large number of preliminary translation jobs, automated translation technology is simple and efficient and can also reduce a significant amount of labor. The world is converging under the catalyst of the Internet, and

communication has become a bridge to globalization, which makes translation technology shine in the language service industry. The semantic problem in the artificial intelligence perspective is one of the most central problems among the many difficulties faced by artificial intelligence. The philosophical debate and questioning of the semantic problem have posed an inevitable and serious challenge to the advancement of AI [3]. It has been sixty-five years since the concept of artificial intelligence (AI) was introduced at the Dartmouth Conference in 1956. Artificial intelligence is a new field that requires the integration of many different disciplines, and it involves cross-disciplinary issues and cutting-edge problems, thus giving rise to a number of "cross-border scholars" with different disciplinary backgrounds, such as Marvin L. Minsky, one of the founders of artificial intelligence and an American mathematician. For

example, Marvin L. Minsky, one of the founders of AI, was not only an expert in computer science but also studied the philosophical issues related to AI; Noam Chomsky, an American philosopher, not only made outstanding contributions to philosophy and linguistics but also thought about the issues related to AI [4]. It is in such a complex disciplinary context that the semantic problem of artificial intelligence emerges in the collision of computer science and philosophical linguistics. Some philosophers have even argued that the semantic problem is a major obstacle limiting the development of AI, and that the inability of syntactic operations to have semantic capabilities has become an inherent shackle, from which AI cannot escape. John Searle, an American philosopher, has suggested that AI can achieve syntactic operations based on symbolic language but cannot achieve understanding and thus does not have semantic capabilities. This idea has not only provoked intense discussions in the field of philosophy but also various responses in the field of artificial intelligence. The semantic problem has long been a complex but important proposition in the philosophical community. Many linguists and philosophers have studied the concepts of “meaning,” “interpretation,” and “intentionality” in depth and have explored the production and realization of semantics in the field of linguistics [5]. The semantic problem of artificial intelligence is a problem that must be faced both in the fields of philosophy and artificial intelligence, both of them testify to the inevitability of the semantic problem from the dimensions of theory and reality respectively, and the problem cannot be solved without the joint efforts of the two fields, that is, the union of theory and reality. There is still a lot of work to be done to solve the semantic problem in concrete terms. In general, the first step is to solve the problem of the conditions and mechanisms of how human beings can realize semantic capabilities, and on this basis, to try to simulate such situations in artificial intelligence systems to demonstrate whether machine languages have the ability to realize semantics. At this stage, neither the philosophical view nor the technical means can solve the semantic problem perfectly yet, and time is needed to explore new ideas [6].

2. Research Background

The translation of Buddhist scriptures has appeared in China since 25 A.D. During the period of Emperor Huan of the Eastern Han Dynasty, as a large number of Buddhist scriptures were introduced to China, the development of translation in China also emerged, and a number of translators such as Xuanzang and Zhiqian emerged, who made outstanding contributions to the development of translation in the world. The second climax of translation in China originated in the late Ming and early Qing dynasties. With the development of economy and transportation and the increasing trade, a large number of western works on natural sciences flowed into the country. Among them, Li Zhizao and Xu Qiguang as representative figures translated a large number of Western scientific and technological works, and some foreign missionaries such as Tang Ruowang and Matteo Ricci made outstanding contributions to the spread

of Chinese culture in the West [7]. The third translation climax was the translation activities before the May Fourth Movement, especially after the defeat of the Sino-Japanese War, which triggered the concern of all people for Western studies. After the First Sino-Japanese War, some rulers thought it was necessary to learn capitalist ideas and to master the long skills of the barbarians to control them. Liang Qichao, Kang Youwei, and Yan Fu were representative figures of this period, especially Yan Fu’s theory of translation with letter, reach, and elegance has been influential to this day. The period from the May Fourth Movement to the founding of New China was a period of unprecedented development and magnificent waves of translation. Since vernacular languages were used in translation, the translations became more popular with the general public and also made the translated texts expand from informational texts to expressive functional texts; the translation methods developed from single semantic translation to communicative translation, and the translation theories were also continuously improved [8]. With the globalization of the world and the continuous intermingling of cultures, translation has ushered in a new climax, and the huge amount of translation work has made human translation unable to meet the translation needs. In the background of this era, machine translation, as well as cloud computing and big data, provides new opportunities for the development of translation. With the increase of information volume, the accuracy of machine translation is also improved, as shown in Figure 1. However, the research on natural language processing and the related application of machine translation in the field of artificial intelligence also face difficulties in cross-domain semantic analysis obstacles such as problems with phrases and syntactic coherence, which are urgent problems for both artificial intelligence and philosophical linguistics, and until the current breakthroughs in computerized symbolic language and cognitive science of thinking, these are still tricky problems for the time being, as shown in Figure 1.

2.1. Review of Foreign Research. Artificial intelligence is an emerging field, but as an interdisciplinary field, there is no lack of research progress [9]. The Turing test in *Computing Machinery and Intelligence* (1950) represents the starting point of AI research. By the end of the 1980s, the symbolist research path encountered a theoretical and practical bottleneck, and AI research entered a period of decline [10]. *What Computers Can’t Do: The Limits of Artificial Intelligence*, published in 1972, and *Mind, Brain, and Program*, published in 1980, were also born in this period. Both had a profound impact on artificial intelligence and semantic issues.

The breakthroughs and advances in the field of artificial intelligence have led to a growing wealth of philosophical research related to artificial intelligence. The next step was that Hogeland, under the influence of his teacher, also took AI as his main research direction and achieved a series of results, mainly introducing the emergence and development of philosophical ideas on AI in recent times and analyzing formalism and computer architecture and discussing

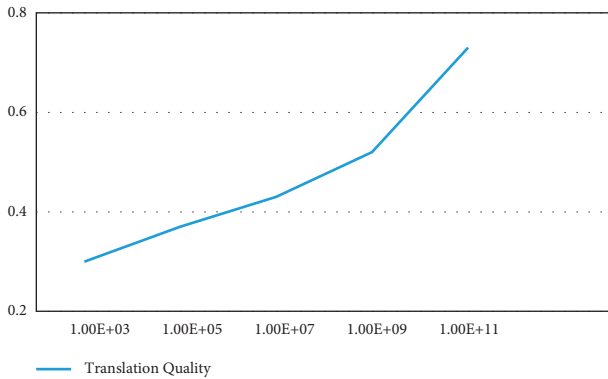


FIGURE 1: The relationship between vocabulary and translation quality.

semantics. In addition, Hogeland compiled the book in 1981 in order to integrate related research in the field of artificial intelligence, and it includes not only research in the field of computer science but also representative articles in the field of philosophy such as Dreyfus' work [11]. The book contains 15 papers on AI and its philosophical issues, and the main arguments in Boden's collection focus on classical judgments of AI and the connectionist approach to research. In 2006, Boden published a monograph, *Mind as Machine: A History of Cognitive Science*, which can be regarded as one of the most comprehensive works in the cognitive science community [12]. The breadth of knowledge and the richness of the material in this work provide a rich ground for the study of artificial intelligence and open up a new field of artificial intelligence research at the level of cognitive science. However, compared with the thinking and direction in the field of computer science, Boden's research on AI is obviously philosophical and critical [13]. According to Yingjin Xu, "We seem to get the impression that Boden is more interested in the entanglements between some AI schools that have become historically silent, but more or less detached from what AI experts are doing at the moment."

In 2011, Floridi, a British expert on the philosophy of information, published a monograph, *The Philosophy of Information*. The book offers a serious and profound exploration of the problem of semantic information. Floridi argues that information is more than a simple physical phenomenon, and that uninterpreted data in a collection cannot be transmitted by encoding and transmission for semantic information. The book gives a realistic dilemma of the semantic information problem; the study of information theory does not require the realization of the transmission of the meaning of information, while the generation and transmission of semantic information are the most important concerns in the field of philosophy.

In addition, there is a journal for the study of artificial intelligence and its philosophy, *Minds and Machines* [14]. Founded in 1991, *Minds and Machines* serves as a platform for communication between the different disciplines of AI research, including philosophy, psychology, cognitive science, and computer science. The preamble issues of the journal also require close cooperation and collaboration

among different disciplines to solve, which has greatly contributed to the development of research in the field of artificial intelligence.

2.2. Review of Domestic Research. After a brief review of foreign research literature, a review of domestic research on this issue is as follows. Due to the division of arts and science education in China, most scholars engaged in philosophical research lack disciplinary backgrounds in mathematics, logic, or even computer science, resulting in not much research focusing on the interdisciplinary field of semantic issues of AI indeed [15]. The domestic research on the semantic problem of AI is mainly divided into two directions: on the one hand, we start from the philosophical problems related to the development of various fields of AI and explore the semantic problem in the process of problem-solving; on the other hand, we start from semantics and conduct research through the path of language logic and language philosophy. The philosophical research on AI in China is still mainly in the form of introduction and translation of foreign research results, but in recent years, there is also a research on the philosophical issues of AI. From the direction of semantics research, both logic and philosophy of language have discussed the semantic problem of AI. Although the research on the semantics of artificial intelligence has not yet taken shape, the discussion and concern on the semantics have never ended [16].

In 2008, Professor Gao Xinmin of Huazhong Normal University published a book entitled "Contemporary Development of Intentionality Theory," which provides a comprehensive analysis of the problem of intentionality in the field of philosophy and discusses the difference and connection between the problem of meaning in different perspectives such as semantics, hermeneutics, and psychology [17]. The semantic properties of individuals are based on relations in the environment, and without informational relations, individuals cannot think and thus do not have semantic properties [18]. In addition to this, "Intentionality and Artificial Intelligence" was published in 2014. The main issues studied in that book are closely related to this study [19]. The problem of intentionality that Intentionality and Artificial Intelligence focuses on is the focus of whether artificial intelligence can be truly realized, and the semantic problem of artificial intelligence that this study tries to solve is also the bottleneck of artificial intelligence [20].

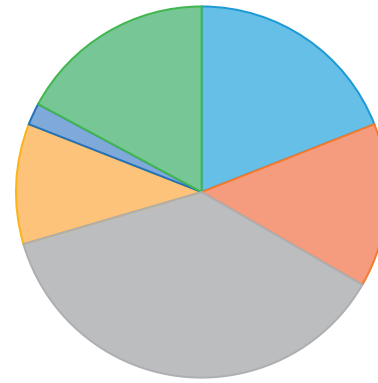
In terms of dissertations, the doctoral dissertation of Weiwei Liu from Shanxi University in 2013, "Research on the Semantics of Science," introduces the relevant theoretical contents of semantics. The dissertation argues that the reason for the complexity and difficulty of semantics research is precise because the research objects and contents of semantics are too much influenced by various schools of thought and positions. The study of semantics is an academic field that requires continuous in-depth research to resolve controversies. In addition to this, Xu Yu's 2016 PhD thesis on "Machine and Language" from the Central Party School explores the issues and controversies raised by language in

the development of artificial intelligence [21]. After sorting out the background of the development of language processing in AI through the dialectical development of both machine and language, the thesis introduces the exploration of the problem of intelligence and language in the field of AI, shows the questioning of the philosophical community in the field of AI through the representative philosophers' views as an example, and finally analyzes the inner logic and the deep reasons for the emergence of the problem of language and intelligence that are difficult to solve [22]. Although the number of articles on the improvement of various problems in translation in combination with big data artificial intelligence and other means has been increasing this year, as shown in Figure 2, according to the literature published in recent years, it can be seen that most researchers mainly explain and elaborate on some problems in machine translation, such as the impact of machine translation on grammar, syntax problems in machine translation, etc., but they report on the syntax problems that do not exist in machine translation. However, they have not followed the trend of using big data technology to analyze the influence of existing artificial intelligence on phrases and syntax in English translation, so the influence of big data technology on phrases and syntax in English translation still needs to be investigated [23], as shown in Figures 2 and 3.

3. Research Ideas

The semantic problem of artificial intelligence requires theoretical knowledge in the field of artificial intelligence technology as well as philosophical analysis due to the complexity and specificity of the disciplines involved [24]. In this study, we try to grasp the causes and development of the semantic problem of artificial intelligence from the intersection of disciplines. On the one hand, it introduces the exploration and attempts of semantic issues in the field of AI, and on the other hand, it introduces the debates and views on semantic issues in the field of philosophy [25].

This study also tries to take this as a clue to combine the technical progress of AI with the philosophical debates to achieve a three-dimensional and comprehensive argument. In this line of thought, it would have been conventional to arrange the materials for alternating arguments between the explorations and debates in the two fields for the semantic problem according to the chronological order. This would have provided a clear contrast between the technical explorations and philosophical theories of the same period. However, in the course of the study, it was found that such a way of thinking jumped around too much, making it impossible to achieve coherence in the content of the study. For the sake of the overall sense and coherence of the study, the research idea has been adjusted. First, we introduce the theoretical background of AI and clarify what the semantic problem of AI is and why it occurs. Finally, possible solutions to the semantic problem are given from the perspective of different disciplinary fields. This arrangement of ideas roughly follows the logical thread of problem emergence and solution [26].



- The effect of machine translation on grammar
- Impact of translation software on text coherence
- Syntactic issues in machine translation
- Impact of translation software on phraseology
- Translation Software on Multiple Meaning Words in Translation
- Problems of Semantics in Machine Translation

FIGURE 2: Percentage of the published content in recent years.

4. Results and Discussion

In the 1950s, mathematical logic and electronic computers were sufficiently mature to be the basis for an important moment in artificial intelligence. First, in 1950 the Turing test asked "Can machines think?" the question of what is artificial intelligence officially opened the research direction; second, the Dartmouth Conference held by John Mc Carthy, Marvin L. Minsky, C.E. Shannon, N. Rochester, and others in 1956. The conference first clarified the concept of artificial intelligence. Since then, AI has had a theoretical foundation and an academic community and has become a new disciplinary field with specific research content and research methods [27]. Of course, for a mature discipline, a research agenda is essential to guide the direction of the discipline's research progress, and this is also true for AI. The disciplinary field of AI has seen the emergence of three different research agendas along the lines of symbolism, connectionism, and behaviorism. For the purpose of the rest of the article, a brief introduction to each of these three lines of thought will be presented.

4.1. Symbolism Based on Symbolic Language. John Hogeland, an American philosopher of artificial intelligence, first proposed the term symbolism. He divided AI into two major categories based on the basis of research and thinking. The first category is directly translated as "old-fashioned classical AI," which refers to symbolic AI [28]. The development of symbolic AI began with Turing machines. However, Turing did not think in terms of Turing machines for AI research, but in terms of "computable numbers." In 1936, Turing wrote the article "Computable Numbers and their Application to the Problem of Determination" published in the Proceedings of the London Mathematical Society. In 1931, the Austrian logician Kurt

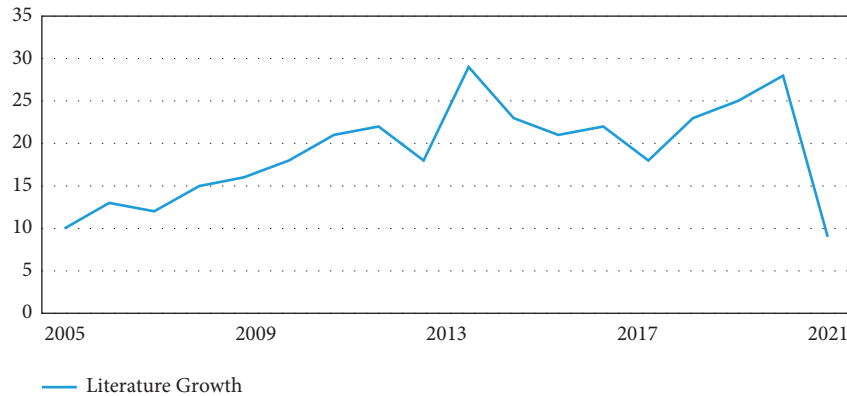


FIGURE 3: Trends of literature growth in recent years.

Gödel proposed the famous “Gödel Incompleteness Theorems.” Since Gödel has shown us that it is axiomatically impossible to prove all reasonable propositions in a system, it is considered inappropriate to think about the relationship between them through machines. Among other things, the rejection of the Judgment Theorems illustrates the current development of mathematical logic, just like the problems of Dividing Angles into Three Equal Angles, Turning Circles into Squares, and Cubic Multiplication, which are bound by the way they are treated. It was on the basis of these elements that von Neumann proposed the principle of stored programs for computers. Finally, it should be emphasized that Turing was not the only one who did not have evidence of the “judgment problem.” In the 1970s and 1980s, neuroscience had already begun to study the function of the brain and related mechanisms and had made a lot of progress. Based on this, Libet tried to use experiments to prove that conscious mental states indicate the illusion in people’s experience [29]. Turing gave his answer to the question of Gödel’s theorem by designing the Turing machine.

The symbolist view is that artificial intelligence is based on symbolic rules of formal language. But other views of different schools of thought point out that the concrete implementation of artificial intelligence requires not only the physical basis of symbolic language but also cannot ignore the important semantic content. The realization of semantic content cannot be achieved by physical symbolic language alone; the ambiguity of language leads to the necessity of external contextual qualifications and denotations. If the generation of intelligence is just some binary sequence of numbers or the design of circuit connections, then human awareness of self-intelligence would be unacceptable.

4.2. Connectionism of Neuronal Distribution Representations.

The article “A logical algorithm for the intrinsic concept of neural activity” published in *Bulletin of Mathematical Biology* in 1943 marked the emergence of connectionist artificial intelligence [30]. Connectionism is based on networks of neurons that use multilinear manipulation and parallel processing of networks to solve representations of mental states.

In 1986, David E. Rumelhart et al. published *Parallel Distributed Processing: Explorations in Cognitive Micro-architecture*, which can be considered as the thriving work of connectionism. Among other things, in their view, the multilinear distributed expression of the associativist mesh-like underlying structure has the following advantages: first, it allows for memory functions. The ability to remember was originally the exclusive domain of the human brain, and the multithreaded distributed representation can process the signals between neurons in parallel, thus simulating the process of fetching memory in the human brain. The second is the extremely strong ability of regulation and adaptation. Different external inputs and stimuli will cause rearrangement among neurons, and this ability to respond to external inputs and adjust its own structure in time is very impressive. Problems that would otherwise require a pre-determined program to solve can be accomplished independently by the connectionist structure on its own.

4.3. Perception-Led Behaviorism. Behaviorism only began to emerge in the 1980s. Behaviorism focuses on the research area of perceiving the external environment and tries to achieve intelligent attempts to guide behavior through perception. The behaviorist view is that artificial intelligence should focus on natural intelligence or humans themselves, and that the path of artificial intelligence should be guided by observing and learning from the patterns of perception-controlled behavior of humans themselves. In 1991, Rodney Brooks of the Massachusetts Institute of Technology (MIT) proposed that there was no need for artificial intelligence to adopt the classical sense-to-model-to-plan-to-act framework, and that the two intermediate steps were unnecessary in his view; only sense-to-act was required.

The emergence of behaviorism and the study of artificial life are inextricably linked. The study of artificial life attempts to trace the essential characteristics of life and along this path to achieve the evolution and transformation of simple to advanced lifeforms. Artificial life mainly relies on genetic algorithms rather than the simulation of intelligence, trying to realize the process of life evolution at the genetic level, and behaviorism is deeply influenced by this. Professor Xiaoli Liu of Renmin University of China summarizes

symbolism, connectionism, and behaviorism as follows: symbolists try to simulate the human brain with symbolic algorithms, connectionists try to construct the brain through parallel computation of artificial neural networks, and behaviorists try to evolve the brain through genetic algorithms, as shown in Figure 4.

Neither linguistics nor semantics has been able to accurately define and grasp semantics, and this is the deep confusion of the semantic problem in the AI perspective. Language serves as a symbolic formal system capable of conveying meaning and thus achieving various other functions. This symbolic formal language system must be able to be endowed with meaning in order to carry the ability to communicate, cognitively experience, or mediate media. The meaning that language has or is given is semantics, which is the essence and foundation of language.

Because of the complexity of the field of research on semantics, the different views of each school and discipline on semantics, and the irreplaceable position of semantics in the progress of research in many fields due to its specificity, there are many different studies on semantics. As the emerging field of artificial intelligence has focused on semantic issues, philosophy, psychology, and cognitive science have also paid attention to semantic issues. In general, the concept of “semantics” that is the focus of this study is the ability to express and even understand the meaning of symbolic languages. The study of semantic issues is not only an important way for humans to understand how they communicate with the outside world but also a doorstop for humans to explore their own thinking processes and cognitive abilities.

In the development of artificial intelligence, there has been a disconnect between syntax and semantics. John Searle was the first to explicitly state that this disconnect is at the heart of the semantic problem. Searle said that “the human mind is not only syntactic, it also has a semantic aspect. Computer programs can never replace the heart for a simple reason: computer programs are only syntactic, while the heart is not only syntactic. The mind is semantic, that is, the mind is not just a formal structure, it has a content.” That is, the syntax of formal language in a computer system is not capable of realizing the semantic content of natural properties, and it is impossible for an artificially intelligent machine to achieve understanding. Is there really a complete disconnection between syntax and semantics? The answer is obviously not exact. Thus, the semantic problem is not unanswerable in the direction of technological development and in the field of philosophical research where new situations may arise. It is just that until the level of science and technology reaches that level of awareness, we have to explore the answers to the questions with an open mind. In fact, from a deconstructive point of view, the theoretical roots of artificial intelligence value logic and use symbolic language, while the requirements of semantic implementation are object-oriented and grounded in reality. Logic and symbols do not exist in reality, but are methods and tools abstracted subjectively by humans. Therefore, solving the semantic problem of AI is equivalent to solving the problem of real objects without abstraction by means of

abstraction-based AI. To solve the semantic problem perfectly, we must reconcile the reactionary connection between foundation and purpose and the contradictory relationship between abstraction and reality.

Along with the development of the information age, artificial intelligence technology has also achieved corresponding results. However, the research on natural language processing, which is the core technology of artificial intelligence, has been difficult because of the difficulty in solving the semantic problem. Therefore, the technical dilemma of AI semantic issues must be sorted out from the semantic barriers faced by natural language processing, mainly in the following aspects.

The first is that the hierarchical structure of language processing implies that a shift from the morphological to the semantic stage of language analysis must be realized. Human analysis and understanding of language are hierarchical processes, which are the consensus of linguistics and computer science for natural language processing research. The process of natural language processing through the human brain can be broken down into two parts: the language is input to the brain, which analyzes and deconstructs the natural language; after that, the brain outputs new results after processing and reconstructing the language. The brain’s analysis and processing of natural language can give the judgment that language can be decomposed into word-level degrees of being reconstructed and deconstructed. Based on this judgment, from a reductionist point of view, it can be argued that natural language can be divided into multiple levels of structure within it, and that computer processing of natural language should follow this hierarchical structure, just like the human brain, to analyze and process language. This requires the computer to simulate as much as possible the analytical logic and grammatical rules that the brain follows when processing language, so that the computer can achieve perfect processing of natural language. In this way, natural language processing can be roughly divided into two modules according to the human brain’s processing of language, namely, the input of language, which requires the computer’s ability to recognize and understand natural language, and the output of language, which requires the computer’s ability to syntactically construct and express natural language. In this process, the recognition and syntactic construction ability of natural language processing need to be realized from the level of vocabulary and sentences, which is the research idea of semantic problem, while the understanding and expression ability of language needs to be realized through the direction of semantic analysis or semantic recognition. On this basis, the study of natural language processing faces the need to cross over from the direction of semantic analysis to the level of semantic analysis. At this stage, the technology and research of semantic analysis have been relatively mature, but the progress in the direction of semantic analysis is still slow, which creates a semantic barrier that natural language processing must face.

Second, neither rationalism based on grammatical rules nor empiricism based on statistical methods can achieve semantic analysis perfectly for the time being. The early

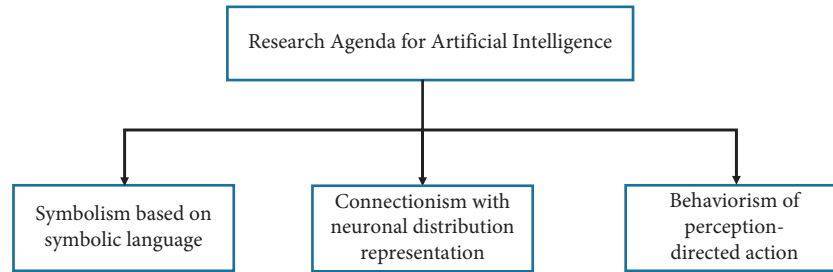


FIGURE 4: Research agenda of artificial intelligence.

research of natural language processing mainly adopts the simulation of human-computer dialogue to realize machine translation. After the emergence of Chomsky's transformational generative grammar, natural language processing realized widespread development and application through semantic analysis and recognition based on Chomsky's. Statistical methods were then added to the mix. At this stage, most of the semantic analysis is based on statistical methods, and the degree and accuracy of the analysis largely depend on the support of data volume. This method has no way to achieve the breakthrough and progress of semantic analysis ability, and it cannot solve the problem of constructing semantic analysis theory. The breakthrough of semantic analysis should be to build a word-level semantic lexicon and to realize the hierarchical analysis structure of the brain for semantics as much as possible; otherwise, it will not be able to break through the bottleneck faced by semantic barriers in theory and practice. With the progress in the field of artificial intelligence, natural language processing is indeed working toward the direction of semantic lexicon. Based on the limitations of statistical methods, natural language processing intends to think differently. It chooses to break away from the reliance on data and chooses to build semantic networks to make a breakthrough from the idea of context analysis and recognition. However, such an idea still has to be limited to syntactic rules and cannot meet the diverse demands on natural language processing. Thus, it seems that the core problem of semantic barriers lies in the fact that there is not a one-to-many logical relationship between syntactic rule-based semantic analysis and complex semantic analysis, but a complex many-to-many conditional relationship, which causes the problem of linguistic ambiguity. Therefore, the construction of a lexical-level semantic lexicon has become an urgent task. Since the 1990s, natural language processing research has indeed made many attempts to build semantic lexicons, but it still cannot escape the shadow of statistical methods and is still limited by the empiricism of the database. Some experts believe that from the perspective of theoretical methods, although the rule-based rationalism method restricts the development of the empiricism-based semantic knowledge base to a certain extent, there are more and more empirical methods that need to be made up by rationalism. Experts also point out that the integration of the two methods is also the current development trend of natural language processing.

Third, the semantic knowledge base based on statistical experience is too subjective and insufficient to support the realization of semantic analysis. The empirical thinking will always have theoretical loopholes, which will cause uncertainty in the results of natural language processing. "The basic semantic framework that constitutes the semantic knowledge base of the frame network starts from the analyst's intuitive judgment, and the establishment of a framework requires some iterative process of recognition. Because of the different knowledge backgrounds between analysts and analysts, and between analysts and users, their ways of thinking cannot be exactly the same, and thus their understanding and awareness of the problem will be different. The resulting frame network is bound to be subjective and uncertain to a certain extent, which cannot be avoided by constructing an empirical semantic knowledge base." Let us take synonymy as an example. The definition and division of synonymy criteria are artificially formulated and involved in computer systems, which makes the language processing at the level of synonymy subjective to humans. It can be seen that the key to the failure of the empirical approach is whether the construction of a semantic lexicon is really suitable to simulate the hierarchical structure of the brain. Not all words and things can be divided into hierarchical categories. In addition to synonyms, there are also things and words with multiple hierarchical properties and category distinctions, and the semantic expression of such things cannot be achieved by a simple hierarchical analysis structure only. On this basis, it is necessary to have the understanding that the capability of the semantic lexicon cannot achieve perfect semantic analysis for the time being, and the evaluation of the capability of the system should be based on its effectiveness and capability in practice.

Finally, the dynamic semantic analysis of the semantic web is difficult to achieve at this stage. The crossing of the semantic barrier of natural language processing cannot be explored only from a one-sided static viewpoint of the solution; after all, language is not just a simple textual expression but also involves the exchange and communication of ideas, which is a dynamic process. Based on this context, Berners-Lee proposed the concept of "semantic web." The semantic web is a kind of semantic Internet based on Internet technology, which can meet the dynamic communication and flexible needs of language processing. This requires the computer's intelligent algorithms and programs to run and be applied openly in the Internet so that the computer can communicate with people without barriers

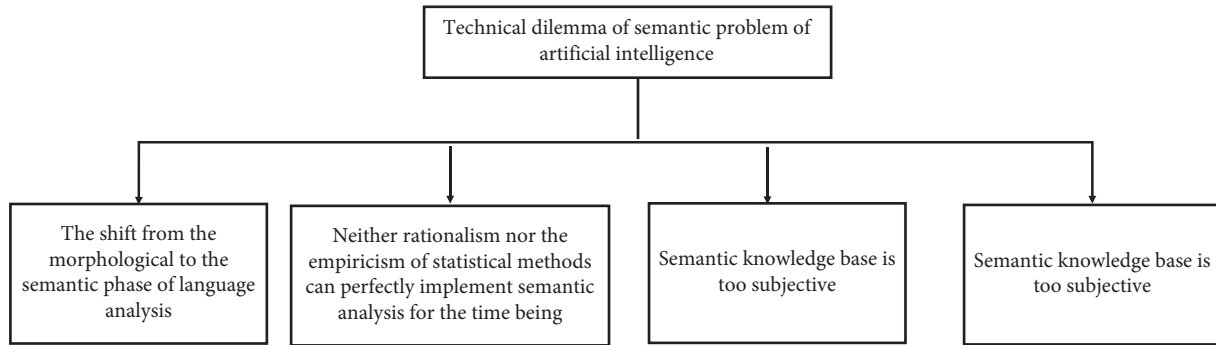


FIGURE 5: The various problems faced by AI semantics at this stage.

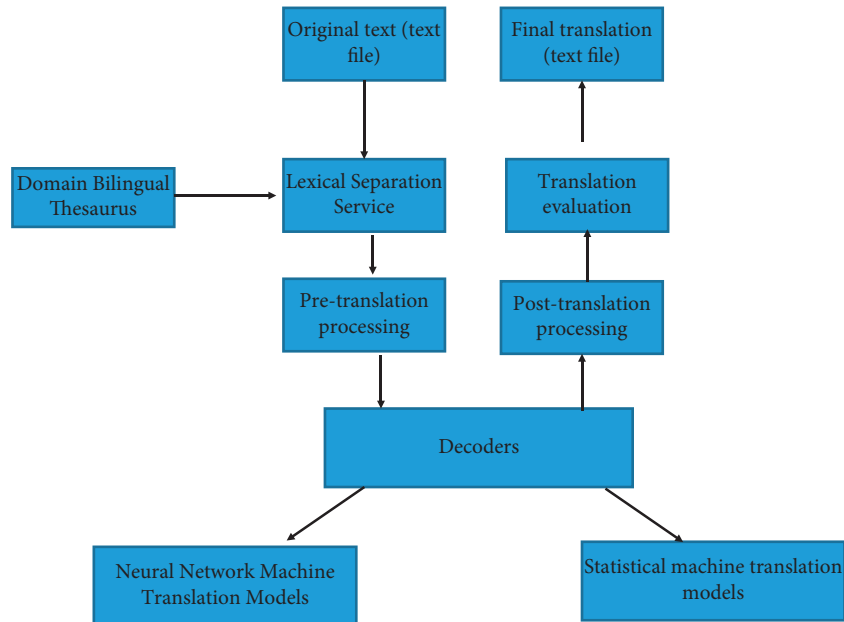


FIGURE 6: Machine translation engine training diagram.

and also so that the computer's processing of language can be continuously learned and improved. This puts new requirements on the computer's natural language processing system, because the instant communication in the Internet is dynamic and evolving and requires the computer to respond and give feedback in time so that the communication can proceed smoothly, as shown in Figure 5.

However, this requires more powerful natural language processing technology that can perform chapter-level semantic analysis, which is still an insurmountable difficulty at this stage, pending new breakthroughs in the technological progress of artificial intelligence. Therefore, the semantic problem in the AI perspective is also an important reason and core motivation to hinder the development of AI breakthroughs, for which many semantic explorations have been made in the technical field to try to solve the problem.

For example, the machine translation engine of New Translation Technology Company is used for training, the basic process is to import 20834 bilingual word pairs into the machine translation engine, and after the machine translation engine learns itself and deep learning, a machine

translation model is generated using "neural network machine translation + statistical machine translation," as shown in Figure 6.

Nowadays, some researchers have also established a series of different framework models to improve the coherence of phrases and syntax in English translation by using artificial intelligence through big data technology; for example, a framework of Chinese to English machine translation system based on the phrase model, as shown in Figure 7.

From the figure, we can also see that the models used are trained: the phrase translation model and the sequencing model are extracted from the parallel corpus with bidirectional word alignment; the language model is trained from the monolingual corpus of the target language. Then, with the system framework diagram of the phrase-based machine translation model, it should be easy to obtain the system framework diagram of the phrase-based interactive machine translation system. What needs to be moved here should be focused on two places: first, the input of the system, in which the interactive case includes not only the source language

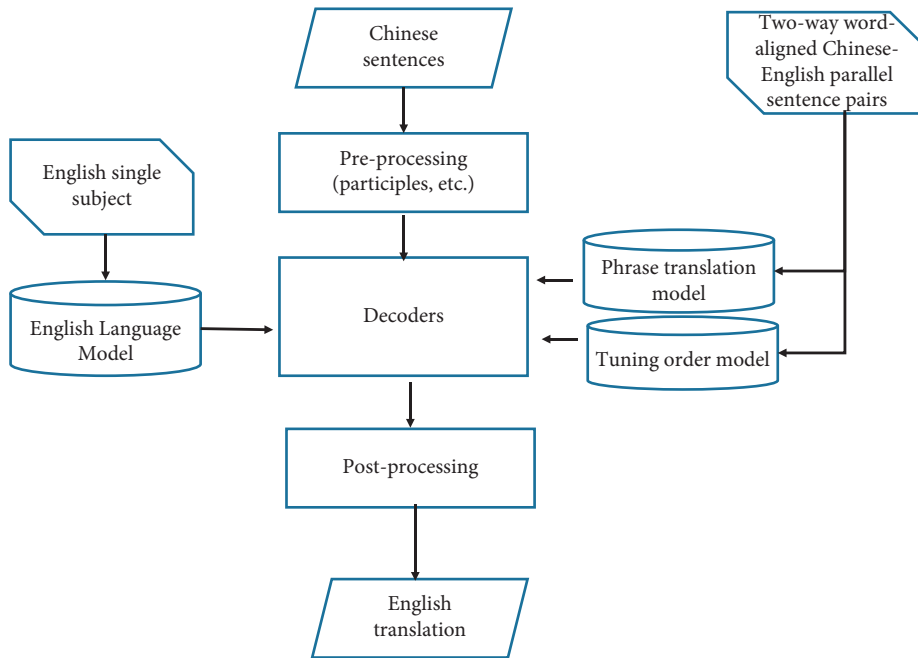


FIGURE 7: Framework of phrase model-based Chinese-to-English machine translation system.

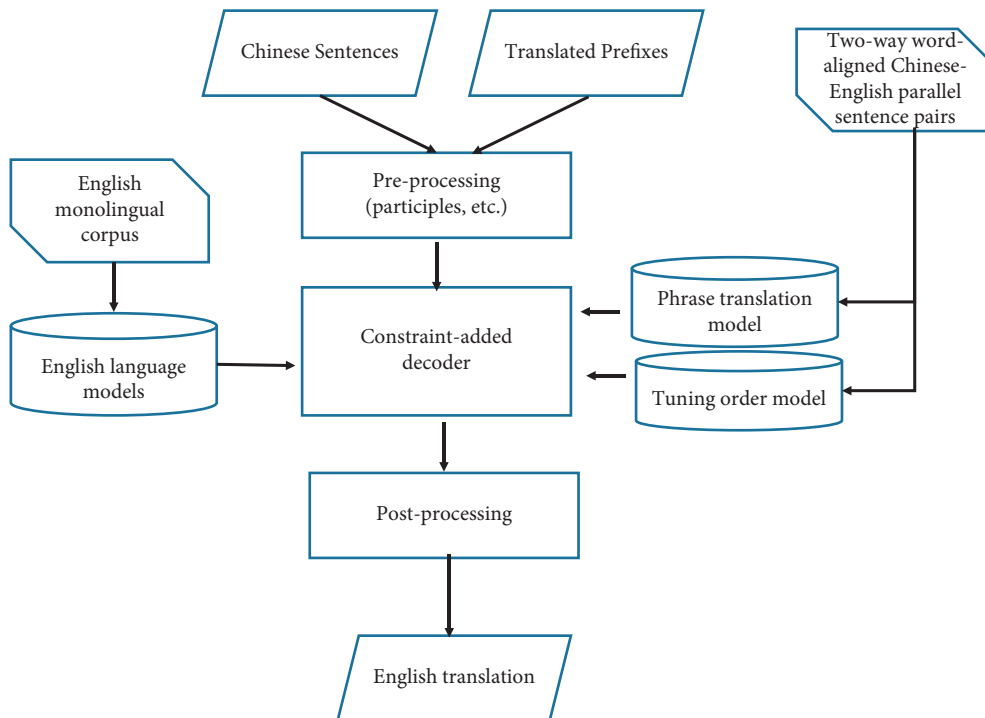


FIGURE 8: Framework diagram of the phrase-based interactive machine translation system.

sentences but also the translation prefixes confirmed by the user; second, the decoder part, where the search and decoding in the interactive environment becomes a restricted decoding process, that is, the paths that do not satisfy the restrictions are not considered. The framework diagram of the phrase-based interactive machine translation

system is outlined, as shown in Figure 8. The framework diagram of the phrase-based model of the interactive machine translation system can be seen to be identical to the framework diagram of the underlying features (models) used in the phrase-based machine translation system, as shown in Figure 8.

5. Conclusion

Today's society has crossed from the information age into the data age, and the basic application based on big data can truly realize the universal and extensive communication mode of language, completely sweep away the difficulties caused by language barriers, and greatly improve the efficiency and quality of translation. As an innovative technology catering to the three elements of translation development, cloud translation further combines machine translation and human through big data, using information technology to bring high speed and rapidity to translation, cooperating with human understanding of the article, combining with context to accurately translate the text, and making it closer to fit the original text. Translators have the creativity and flexibility that cannot be replaced by machine translation, and the combination of cloud technology and human is the optimal solution for today's translation business. For example, under the mode of crowdsourcing translation, a large number of translation teams and volunteers from the network participate in translation, which better solves the difficulties brought by heavy translation tasks to translators, thus realizing a large number of translation results output. To sum up, the continuous development of big data and cloud computing technology makes translation technology become more and more mature, and the complementary translation of efficient and fast machine translation and accurate human translation to fit the original text will become the new normal. In translation practice, neither the efficiency and convenience brought by machine translation to people in the translation business nor the translator's precise positioning of the translated text should be neglected; while the cloud technology should be fully utilized to pool translation resources, the translator's overall grasp of the text is needed to avoid unnecessary phrases and syntactic errors.

Data Availability

The dataset can be accessed upon request to the author.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

This work received Jilin Provincial Educational Science Planned Research Project: Research on the Development of Autonomous Learning of English Majors in Private Universities Based on SPOC Blended Teaching Mode, under GH20475.

References

- [1] H. Huang, "The cooperative principle based study on passive sentence translation in science and technology English—to improve the quality and efficiency of non literature translation [J]," *Open Journal of Modern Linguistics*, vol. 12, no. 2, 2022.
- [2] T. Du, "Characteristics and translation strategies of petroleum science and technology English text[J]," *Frontiers in Economics and Management*, vol. 3, no. 2, 2022.
- [3] J. Zhang and X. Ning, "Data-driven teaching model design of college English translation using intelligent processing technology[J]," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 6559772, 2022.
- [4] A. Syed Abdul Basit, Wahid Abdul, and G. S. Kumar, "Machine translation system using deep learning for English to Urdu[J]," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.
- [5] J. Yang and K. Vinoth Babu, "Digital mining algorithm of English translation course information based on digital twin technology[J]," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 9741948, 2021.
- [6] M. Kou, "Analysis on the innovative approaches of college English translation teaching based on new media technology [J]," *Learning & Education*, vol. 10, no. 3, 2021.
- [7] Y. Dong, "The studies on the translation strategies of the scientific English[J]," *International Journal of Educational Technology*, vol. 2, no. 3, 2021.
- [8] X. Shi, "Ocean surface temperature simulation based on big data and accuracy of English translation of foreign trade vocabulary[J]," *Arabian Journal of Geosciences*, vol. 14, no. 17, 2021.
- [9] T. P. Shu and L. H. Chang, "Students' progressive behavioral learning patterns in using machine translation systems – a structural equation modeling analysis[J]," *System*, 2021.
- [10] C. Liang, "Acquisition of English translation examples incorporating shallow syntactic analysis based on network technology[J]," *Journal of Physics: Conference Series*, vol. 1992, no. 2, 2021.
- [11] L. Nie, "Research on the translation system of English cross-cultural communication based on computer technology[J]," *Journal of Physics: Conference Series*, vol. 1992, no. 3, 2021.
- [12] Y. Wang, "Evaluation of English translation accuracy of green plant surface irrigation and food words based on image processing[J]," *Arabian Journal of Geosciences*, vol. 14, no. 15, 2021.
- [13] M. Chen, "An investigation on the application of AI technology in college English teaching[J]," *International Journal of Educational Management*, vol. 6, no. 2, 2021.
- [14] X. Meng, "On the application of artificial intelligence in English and Chinese Translation Teaching[J]," *International Journal of Social Sciences in Universities*, vol. 4, no. 2, 2021.
- [15] M. Organ, "Translation technology in English studies within the system of higher education in Poland[J]," *International Journal of English Linguistics*, vol. 11, no. 4, 2021.
- [16] F. Li, "The application of Chinese-English translation skills in English writing[J]," in *Proceedings of the 2021 5th International Conference on Informatization in Education, Management and Business*, vol. 22, Beijing, China, June 2021.
- [17] L. Wang, "Urban land ecological evaluation and English translation model optimization based on machine learning [J]," *Arabian Journal of Geosciences*, vol. 14, no. 11, 2021.
- [18] Ye. Jin, "On the innovation of college English translation teaching model under multimedia technology[J]," *Journal of Physics: Conference Series*, vol. 1915, no. 4, 2021.
- [19] R. Li, "Reflections on college English translation teaching mode based on network technology[J]," *Journal of Physics: Conference Series*, vol. 1915, no. 2, 2021.
- [20] H. Wu, "Multimedia interaction-based computer-aided translation technology in applied English teaching[J]," *Mobile*

- Information Systems*, vol. 2021, Article ID 5578476, 10 pages, 2021.
- [21] Z. Wu, "Research on English translation based on computer technology[J]," *Journal of Physics: Conference Series*, vol. 1744, no. 4, 2021.
- [22] Z. Zhao, "Research on English translation skills and problems by using computer technology[J]," *Journal of Physics: Conference Series*, vol. 1744, no. 4, 2021.
- [23] Z. Hou, "The metaphor processing and translation skills of computer technology in English language translation[J]," *Journal of Physics: Conference Series*, vol. 1744, no. 4, 2021.
- [24] F. Zalinda, H. Budi, and M. D. Dia an, "Analysis of translation techniques in thesis abstracts of English and Indonesian language and literature undergraduate students[J]," in *Proceedings of the Thirteenth Conference on Applied Linguistics (CONAPLIN 2020)*, p. 546, Amsterdam, Noord-Holland, The Netherlands, November 2021.
- [25] L. Wang, Xu. zhou, and Y. Jiang, "The application of intelligent translation technology in English translation of space science and technology[J]," in *Proceedings of the 2021 International Conference on Public Relations and Social Sciences (ICPRSS 2021)*, p. 586, Kunming, China, October 2021.
- [26] R. Hu, "Functional skopos theory of science and technology English translation[J]," in *Proceedings of the 2020 2nd International Conference on Humanities, Cultures, Arts and Design*, vol. 15, Sydney, Australia, January 2020.
- [27] Li. Han, "Technical English translation techniques and strategies[J]," in *Proceedings of the 2020 2nd International Conference on Humanities, Cultures, Arts and Design*, vol. 15, Sydney, Australia, January 2020.
- [28] Li. Han, "Transforming the cross-cultural perspective and analysis of translation techniques in English translation[J]," in *Proceedings of the 2020 2nd Asia-Pacific Conference on Advance in Education, Learning and Teaching*, p. 66, Guangzhou, China, November 2020.
- [29] R. Wang and M. Dong, "Aesthetics in English for science and technology based on the translation aesthetics[J]," *Open Access Library Journal*, vol. 07, no. 11, 2020.
- [30] W. Qu, "Research on strategies of modern information technology assisting children's English teaching[J]," in *Proceedings of the 2020 8th International Education, Economics, Social Science, Arts, Sports and Management Engineering Conference*, vol. 64, New Delhi, India, November 2020.