

## Research Article

# English Translation Recognition Model Based on Optimized GLR Algorithm

Zuo Guangming 

*Huaiying Institute of Technology, Huian, Jiangsu 223000, China*

Correspondence should be addressed to Zuo Guangming; [zuoguangming@hyit.edu.cn](mailto:zuoguangming@hyit.edu.cn)

Received 7 April 2022; Revised 24 April 2022; Accepted 23 May 2022; Published 15 June 2022

Academic Editor: Jie Liu

Copyright © 2022 Zuo Guangming. 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 current era of information technology, people's requirements for English translation are gradually increasing, and the need for a computer to understand and translate English language is becoming more urgent. In order to accurately identify phrases, this paper proposes an English translation recognition model based on optimized GLR algorithm, which can improve the accuracy of recognition by locating phrases in the text.

## 1. Introduction

Translatology is an important theoretical basis of language translation in China. English translation activity class under the background of artificial intelligence can provide an important boost to the development of translation studies. To ensure great achievements in Chinese English translation, artificial intelligence translation technology should be improved. In the past, it was mainly through the training of a large number of topics to let students explore the rules of translation and thus accumulate translation experience. Traditional translation tools are mostly paper dictionaries. Now, in the context of artificial intelligence, artificial intelligence translation technology can be upgraded to make it fit the concept of translation teaching to ensure that English translation can be changed with the help of new artificial intelligence translation technology.

With the rapid development of economy, the internet industry is developing rapidly, and the status of English translation in world trade is gradually improving. Machine translation technology can overcome many problems in human translation and reduce the economic consumption and time consumption of human translation [1–3]. In the current era of information technology, people's requirements for English translation are gradually increasing, and the need for a computer to understand and translate English language

is becoming more urgent [4–8]. The English translation ability of the computer directly affects the translation result, however, there will be some grammatical errors in the translation result, which will cause problems in the translation result and affect the final decision of English translation. Therefore, in the past studies, many experts have proposed automatic recognition methods for machine English translation errors, thus minimizing the errors in English translation [9, 10].

With the continuous development of global economic integration and the deepening of world trade, the contacts between countries are also constantly deepening, and the frequency of personnel exchanges between countries also increases accordingly. Language is a unique function of human beings and the main means of human communication. Because of the different language environment in each country, language barriers greatly hinder the communication between different countries. Translation robots are born to break down language barriers and enhance communication between countries. The translation robot is mainly composed of speech input and output system, language processing system, and language translation software. Through the combination of multiple software and hardware, it forms a translation platform that can understand multiple languages. The translation robot stores a large amount of language information inside and has intelligent

functions, such as automatic learning, analysis, and memory, which can help humans translate various languages and simplify the communication process between people from different countries. Nowadays, English is the most frequently used language in the world, and English translation robot has the most application scenarios and a wider range of applications. With the continuous upgrading and development of English translation robots, although they can effectively translate a variety of languages into English, translation errors are easy to occur because of the inflexible translation content, which affects people's communication. How to automatically detect the translation errors of English translation robots is one of the most urgent problems in the field of translation robots.

Because of the current machine translation results, there are certain problems. After using the server to compare the full text, the grammar and rules of each language can be obtained. It can be found that the machine translation has low accuracy and low efficiency. Therefore, we should use a more intelligent piece of technology for machine translation [11–15]. In the actual testing process of machine translation products, such as Baidu and Google translation software, the quality of translation results is quite different from that of actual professional manual translation. The current machine translation technology cannot meet the requirements, and the market urgently needs a high-performance [16, 17]. Because of the development of artificial intelligence, many researchers have sought to help with translation work through computer-aided translation (CAT). The central idea of CAT is that the translation results are usually taken as auxiliary reference, and the user usually judges the quality of the translation and then makes a choice. In addition, through the use of corpus, the vocabulary of all industries can be sorted out, more in line with the actual needs of users. The correct use of frequently translated words can greatly reduce the amount of repeated translation work and greatly improve the accuracy of translation [18–20].

For an English translation robot, translation accuracy is the main evaluation index of robot application performance. Because of the influence of internal storage information and external environment, the frequency of translation errors is high, which is not conducive to the development and application of translation robots. According to the existing research results, the existing translation error detection system cannot detect translation errors effectively because of the defects of hardware and software. Zhang et al. used the neural machine translation method to predict the Chinese and English translation results and completed the identification of translation errors in the process of prediction. Huang Dengxian compares phrase words and phrase corpus to analyze part-of-speech and syntax. The author further obtains the English syntactic structure that needs to be translated, and the errors are gradually transmitted and accumulated, which eventually leads to the disadvantage of low translation accuracy. Then, the author designs vocabulary semantics based on HowNet similarity and the logarithmic linear model, saves the corresponding bilingual

corpus in the form of Chinese-English dependency tree to string, provides structured processing of language dependencies, ensures the corresponding relationship between Chinese and English, and calculates the operation input of HowNet that needs to translate sentences with the same example. The semantic similarity of words in the source language of the library further improves the accuracy of translation, and the translation results have high accuracy [21].

Through the summary of the above literature, it is found that intelligent phrase recognition is an important step of speech recognition, and its principle is to realize automatic translation and combination by analyzing its part of speech and syntax, and output the results [22–24]. In the field of machine translation, intelligent phrase recognition is the key technology that can satisfy the selection of translation samples and the accurate alignment of parallel corpus. The technology of intelligent phrase recognition can effectively reduce grammatical ambiguity. The focus and difficulty of current English translation is structural ambiguity. Based on the GLR model in machine translation, this paper analyzes the structural ambiguity in some phrases through the syntactic function of the model, so as to facilitate the understanding of the entire semantics, solve the problems existing in the current English translation, and improve the efficiency and accuracy of the entire translation.

## 2. GLR Algorithm

*2.1. Traditional GLR Algorithm.* GLR algorithm is an extended LR analysis algorithm. The introduction of graph stack and analysis forest can effectively solve the ambiguity problem that an LR algorithm cannot handle, and its analysis speed is fast, which has great advantages in simple syntactic analysis.

In this paper, the GLR algorithm is used to identify and analyze the phrases in each fragment. The GLR algorithm is based on an extended context-free grammar, which is a five-element formula  $G = (V_T, V_N, V_F, P, S)$ , where  $V_T$  is a nonempty finite terminal symbol set,  $V_N$  is a nonempty finite nonterminal symbol set, and the intersection of  $V_T$  and  $V_N$  is empty.  $V_F$  is a constraint function set, which is a nonempty finite set that can be reduced by production only when the conditions are satisfied.  $P$  is the generation formula set, and  $P \rightarrow \langle D, T, M \rangle$ ,  $D \in (V_T \cup V_N)^+$ ,  $D$  is the right-hand symbol string of the production.  $T \in V_T \cup V_N$ ,  $T$  is the central symbol of the production,  $M \in V_F$ ,  $V_F$  detects the part-of-speech and semantic features of  $T$ . When the symbol string at the top of the stack can be reduced to  $P$ , specify its central symbol as  $T$ .  $S$  is the starting symbol set,  $S \in V_N$ .

The steps of GLR algorithm analysis are as follows:

- (1) Initialization. State  $O$  is pushed onto the stack. The analysis pointer points to the input symbol to be analyzed, and the termination flag is cleared.
- (2) Symbol mapping. If there is no end flag, the current input symbol is mapped to the analysis table terminator using a mapping function.

- (3) Check the ACTION table to determine the operation that will be performed next.
  - ① If it is moved up, the current state and current symbol will be pushed, and the analysis pointer will be moved down.
  - ② If it is a statute, The constraint function checks whether the conditions are met. If the conditions are met, the center word pointer points to the corresponding center word. If not, the end mark is marked.
  - ③ If it is terminated, it refers to the pointer to the analysis table terminator “error” the current input character is remapped to the analysis table terminator to continue analysis, and then set the end flag.
  - ④ If it is accepted, the recognizable phrase completes the analysis, pops up the syntax tree at the top of the symbol stack, and returns.
  - ⑤ If it is an error, it refers to the “error” for the terminator of the analysis table, which belongs to the analysis failure, restores the initial state, and returns.
- (4) Continue to execute the next action in sequence until the end of the analysis.

*2.2. Improved GLR Algorithm.* In general, the GLR algorithm is still unable to meet the existing accuracy because of its high probability of coincidence in the results. In this paper, the classical GLR algorithm is improved, and the phrase center is proposed to analyze the phrase structure. The improved GLR algorithm realizes the likelihood calculation of the prefixes and postfixes of phrases by means of quantization, as shown in formula (1).

$$G_E = (V_N, V_T, S, \alpha). \quad (1)$$

In formula (1),  $V_N$  represents the cyclic symbol cluster,  $V_N \neq \emptyset$ .  $V_T$  represents the termination symbol cluster.  $V_T \neq \emptyset$ , and the elements in  $V_T$  and  $V_N$  do not overlap.  $S$  represents the start symbol cluster, which is an element in  $V_N$ .  $\alpha$  represents phrase action clusters.

Assuming that  $P$  is any action in  $\alpha$  and  $P$  exists in  $V_N$ , formula (2) can be obtained by derivation.

$$P \longrightarrow \{\theta, c, x, \delta\}. \quad (2)$$

In (2),  $\theta$  represents the right side of the action,  $C$  represents the center point,  $x$  represents the constraint value, and  $\delta$  represents the marking mode.  $\theta$  and  $C$  are located in both  $V_T$  and  $V_N$ , and  $\delta$  can be located in  $V_T$  or  $V_N$ .

### 3. English Intelligent Recognition Algorithm

*3.1. Create Phrase Corpus.* Corpus is mainly used to store phrases, which can accurately mark the parts of speech in English, further standardize the function of phrases, and make machine translation more accurate.

There are more than 700,000 words in the corpus of the intelligent recognition model constructed in this paper,

which can meet the actual demand. In this paper, we distinguish English and Chinese phrase corpora by distinguishing the tenses of phrase corpora. The marking process is divided into layers, data, and processing, and the processing adopts the man-machine active communication mode to carry out the operation of English translation.

*3.2. Phrase Corpus Part-of-Speech Recognition.* The dependency relationship of phrases is analyzed using syntax, and the establishment of syntax tree is realized. The part-of-speech recognition of phrases is a key step in the intelligent recognition algorithm of machine translation, which can deal with the grammatical ambiguity of a large number of sentences, phrases, and words. Each sentence is divided into several words into English sentences, and the processed words are aligned to form phrases. Meanwhile, the parts of speech of the words are marked out by judging the context of the sentence. Finally, the syntactic tree of sentences is formed by analyzing the dependency of phrases. Through this method, the timeliness and accuracy of machine translation can be improved, and the processing capacity of phrase corpus can be significantly increased. GLR is a commonly used algorithm in part-of-speech recognition, which is mainly used to judge the contextual relationship of phrases. Its core theory is based on the dynamic recognition of forms and unconditional transfer statements.

In traditional GLR, the execution of each step is carried out through operation shift instruction and reduction instruction. In this process, the beginning and end of each operation are carried out according to specific standards. In the process of translation, if grammatical ambiguity is detected, it is necessary to use the geometric structure linear table of syntactic analysis to call up the analytic linear table, expand and identify the content of the phrase, select the optimal content, and transfer it to different recognition channels for recognition.

*3.3. Correction Process of Phrase Intelligent Recognition Algorithm.* In current machine translation algorithms, the matching results of segmented phrases and phrase corpora are often regarded as the final machine translation results, which lack the analysis of the context in which the phrases are located, and excessively rely on the part of speech analysis of phrase corpora, resulting in inaccurate final translation results. Therefore, this paper further considers to correct the results of part-of-speech analysis. In the process of part-of-speech analysis and correction for the improved GLR algorithm, in view of the error points in the part-of-speech recognition results of phrases using analytic linear tables, the correction process is carried out by checking the tagged content in the corpus, as shown in Figure 1.

However, the reduction expresses that the previous constraints have no effect or that there is a problem in the loop process, and it is necessary to clarify the syntactic function to identify the constraints again. The advance indicates that there is no structural ambiguity in the ongoing syntactic function recognition, and the phrase part-of-speech recognition result is accurate. At this time, the

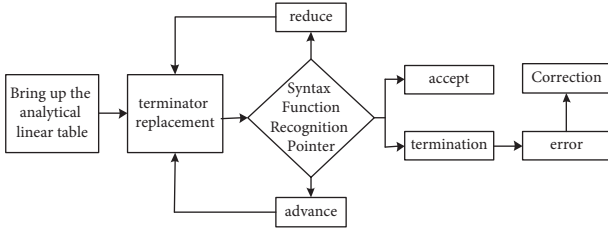


FIGURE 1: Calibration flow chart of intelligent recognition algorithm.

acceptance pointer should be selected for use. The receive pointer and the advance pointer usually appear together. If this condition is not met in the process and only a certain pointer appears, it means that there is an error in the loop or an error in the algorithm. Then, the analytical linear table needs to be called up again, and the part-of-speech recognition results that have been agreed by default before are withdrawn.

During the operation of the improved GLR algorithm, the type of pointer should be identified before the termination is replaced. If it is a protocol pointer, the constraint conditions of pointer should be detected in the phrase corpus. If it does not, it goes directly to the termination pointer.

#### 4. English Translation Intelligent Recognition Model

The functions of the intelligent recognition model for English translation are designed. The received voice signal is obtained through the data acquisition device, and then the English signal is input to the processing system. The data signal is processed, the results are output in the display, and the user can view the automatic identification results of English translation through the display or the client.

**4.1. English Signal Processing.** Detailed design is required after model design, and English signals are collected and processed in a planned way. However, because of the interference factors of speech signals, the collected speech signals should be processed to improve accuracy. Figure 2 shows the processing process of English signals.

The digital filter is used for signal weighting processing, and the stress detection system is improved. Firstly, F1 is used to represent the first formant of vowel spectrum characteristics, and F2 is the second formant. Using the classifier to output confidence, the vowel intonation is obtained, and the best speech signal is selected. The calculation formula of weighting signal  $Y(n)$  is as follows:

$$y(n) = T[x(n)] = ax(n) + b. \quad (3)$$

To make the analysis result of speech signal more accurate, the speech signal is divided into  $T$  frames.

$$z(n) = \frac{1}{t} y(n). \quad (4)$$

To clearly display the speech effect, select the rectangular window  $W(n)$ .

$$W(n) = w(n) \times z(n). \quad (5)$$

**4.2. Extract Feature Parameters.** To further improve the operation efficiency of the system and reduce the data interference unrelated to the voice signal, it is necessary to unify the relevant information data to find the parameter characteristics and then realize the subsequent calculation. Figure 3 shows the structure of extracted feature parameters.

The continuous spectrum of aperiodic continuous time signal is calculated by Fourier transform, however, the discrete sampling value of continuous signal is obtained in the actual control system. Hence, the signal spectrum is calculated by discrete sampling value. A finite length discrete speech signal is improved to obtain the following formula:

$$X[K] = \sum_{n=0}^{N-1} x[n]e^{-j2\pi/Nnk}, \quad k = 0, 1, 2, \dots, N. \quad (6)$$

Convert a discrete speech sequence to a Mel frequency scale.

$$\text{Mel}(f) = 2579 \lg\left(1 + \frac{f}{700}\right). \quad (7)$$

Through DTC calculation of the output filtering, the characteristic parameter  $P$  of the speech signal  $W(n)$  is obtained.

$$P = \overline{Z}_{n=1}^N F(l)w(n)\cos(\pi n(M + 0.5)). \quad (8)$$

After the spectrum of a speech signal is generated, it is processed by weighting, windowing, and framing. Each short-time analysis window can get spectrum information through fast Fourier transform. Then, use the Mel filter to get an MFCC two-dimensional graph.

Using the above method, related speech signal parameters are extracted from rhythm, speed, pitch, and intonation.

#### 5. Experiment Analysis

In this paper, three machine translators were used to translate 50 phrases and 50 network random sentences. English-Chinese translation professionals also translated the above sentences. Graders scored the results of machine translation by comparing the results of machine translation.

As can be seen from Figure 4, the machine translation results of the proposed algorithm are optimal compared with other algorithms in terms of recognition accuracy, speed, and updating ability. As can be seen from Figure 5, the improved GLR algorithm in this paper has the highest score, while the statistical algorithm has the lowest score.

The comparison experiment in this paper also adopts the experiment on actual translation cases, and the sentence ‘‘Xi’an Price Bureau on beef noodle Price Limit’’ is selected for

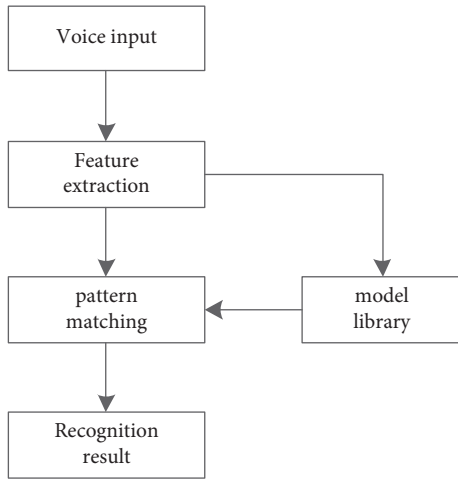


FIGURE 2: Processing of English signals.

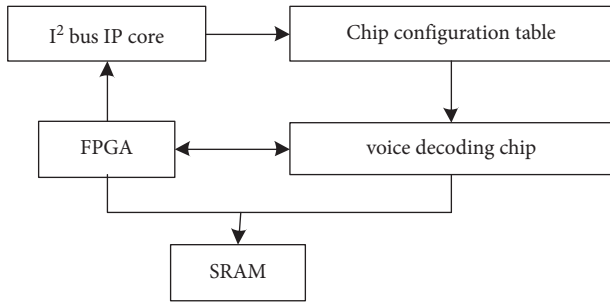


FIGURE 3: Feature parameter extraction map.

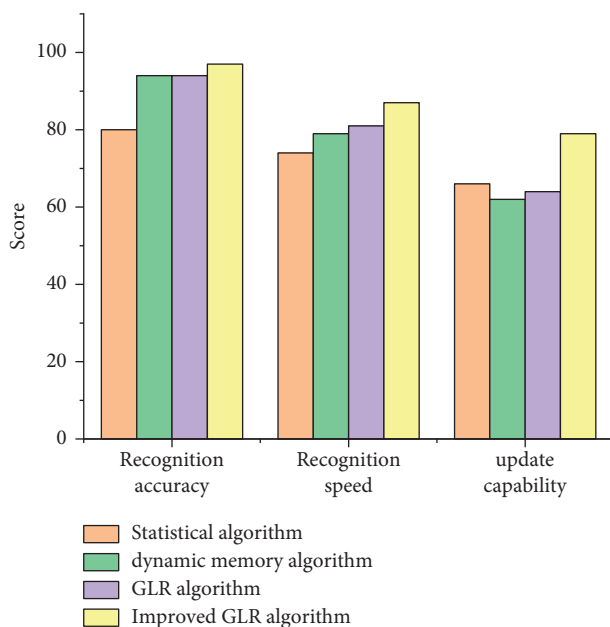


FIGURE 4: Evaluation results of four algorithms.

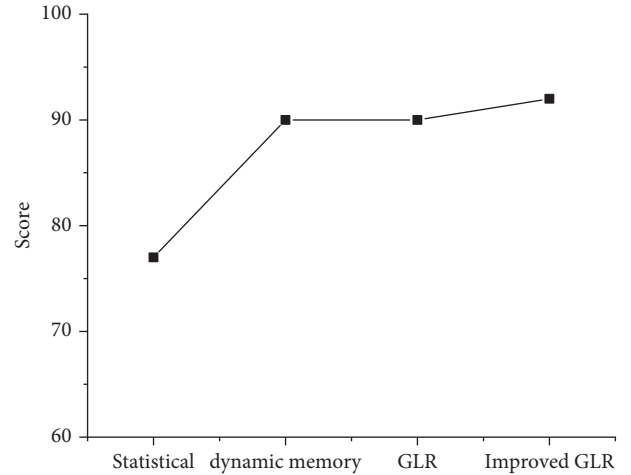


FIGURE 5: Comparison of scores of four algorithms in comprehensive test.

TABLE 1: Comparison of translation example results.

Translation method	Translate content
Improved GLR algorithm	Xi'an price bureau gives the explanations of beef noodles reduce: only because of the excessive price raises.
Human translation	Xi'an price bureau gives the explanations of price control on beef noodles: it is only because the raises have been too large.

translation. The experimental comparison results of machine translation and human translation based on statistical algorithm, dynamic memory algorithm, and improved GLR algorithm are shown in Table 1.

As can be seen from Table 1, compared with other algorithms, the algorithm in this paper is more accurate, and the recognition accuracy reaches more than 95%, reaching the same level as human translation, indicating the efficiency and feasibility of the improved GLR algorithm in machine translation.

## 6. Conclusion

With the rapid development of economy, the internet industry is developing rapidly, and the status of English translation in world trade is gradually improving. Machine translation technology can overcome many problems in human translation and reduce the economic consumption and time consumption of human translation. In the current era of information technology, people's requirements for English translation are gradually increasing, and the need for a computer to understand and translate English language is becoming more urgent. Computer's English translation ability has a direct impact on the application effect of translation results. In this paper, using the generalized maximum likelihood ratio detection algorithm based on improved machine translation, set the phrase corpus using

this algorithm, the library's size to 740,000 English words, and by constructing the phrase structure through the central phrase and calibrating the structural ambiguity according to the syntactic function, the content of recognition can be obtained, and the actual position range of phrases in translation can be determined, so as to solve the problems existing in current English translation and improve the accuracy and efficiency of recognition.

## Data Availability

The dataset can be accessed upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] J. L. Yu, X. Ma, and Z. Qu, "English translation model based on intelligent recognition and deep learning," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 3079775, 9 pages, 2022.
- [2] Y. Chen and S. H. Ahmed, "Business English translation model based on BP neural network optimized by genetic algorithm," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 2837584, 10 pages, 2021.
- [3] Z. Li, "The construction of the turning classroom of business English translation teaching in higher vocational education under the Internet + environment," *Frontiers in Educational Research*, vol. 2, no. 5, 2019.
- [4] Y. Ruan and N. Jan, "Design of intelligent recognition English translation model based on deep learning," *Journal of Mathematics*, vol. 2022, Article ID 5029770, 14 pages, 2022.
- [5] F. Fang, "The construction of college English translation teaching mode in the information technology environment," *International Journal of Social Sciences in Universities*, vol. 3, no. 3, 2020.
- [6] P. Zhou, "Comparative study of traditional teaching method and ISAS teaching method in business English translation teaching in colleges and universities," *Frontiers in Educational Research*, vol. 2, no. 3, 2019.
- [7] X. Wang, "Translation correction of English phrases based on optimized GLR algorithm," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 868–880, 2021.
- [8] M. Z. Sheriff, M. N. Karim, H. N. Nounou, and M. N. Nounou, "Process monitoring using PCA-based GLR methods: a comparative study," *Journal of Computational Science*, vol. 27, pp. 227–246, 2018.
- [9] Y. Guo, D. Zhou, J. Cao, R. Nie, X. Ruan, and Y. Liu, "Gated residual neural networks with self-normalization for translation initiation site recognition," *Knowledge-Based Systems*, vol. 237, 2022.
- [10] C. Rajesh Kumar, J. Bansal, and P. Bansal, "Machine translation model for effective translation of Hindi poetries into English," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 34, no. 1, 2022.
- [11] T. Tian, S. Chai, T. Jin, and H. Huang, "A French-to-English machine translation model using transformer network," *Procedia Computer Science*, vol. 199, 2022.
- [12] L. Fei and J. Su, "The construction of machine translation model and its application in English grammar error detection," *Security and Communication Networks*, vol. 2021, Article ID 2731914, 11 pages, 2021.
- [13] Y. Zhang and X. Ning, "A study on the intelligent translation model for English incorporating neural network migration learning," *Wireless Communications and Mobile Computing*, vol. 2021, pp. 1–10, Article ID 1244389, 2021.
- [14] L. Wang, "Retraction Note to: u," *Arabian Journal of Geosciences*, vol. 14, no. 22, p. 2430, 2021.
- [15] Ke Sun, T. Qian, C. Xu, and M. Zhong, "Context-aware seq2seq translation model for sequential recommendation," *Information Sciences*, vol. 581, 2021.
- [16] L. H. Baniata, I. K. E. Ampomah, and S. Park, "A transformer-based neural machine translation model for Arabic dialects that utilizes subword units," *Sensors*, vol. 21, no. 19, p. 6509, 2021.
- [17] T. Yang, S. Zhao, He Chen, and T. Liu, "A neural machine translation model based on sequence to dependency," *Journal of Physics: Conference Series*, vol. 2030, no. 1, 2021.
- [18] J. Wu, B. Yang, and L. Tian, "Cross language relation extraction method based on attention of bidirectional translation model," *Journal of Research in Science and Engineering*, vol. 3, no. 7, 2021.
- [19] X. Li and H. Xing, "English machine translation model based on artificial intelligence," *Journal of Physics: Conference Series*, vol. 1982, no. 1, 2021.
- [20] H. Zheng, "Research on computer intelligent proofreading system of improved English phrase translation model," *Journal of Physics: Conference Series*, vol. 1871, no. 1, 2021.
- [21] N. S. Khan, A. Abid, and K. Abid, "A novel natural language processing (NLP)-Based machine translation model for English to Pakistan sign language translation," *Cognitive Computation*, vol. 12, 2020.
- [22] Y. Xia, "Machine translation; report summarizes machine translation study findings from chongqing three gorges university research on statistical machine translation model based on deep neural network," *Computers, Networks & Communications*, vol. 102, 2020.
- [23] J. Su, J. Zeng, D. Xiong, Y. Liu, M. Wang, and J. Xie, "A hierarchy-to-sequence attentional neural machine translation model," *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, vol. 26, no. 3, 2018.
- [24] J. Su, Z. Wang, Q. Wu, J. Yao, F. Long, and H. Zhang, "A ttriggered translation model for statistical machine translation soochow University, Automation department, xiamen university. A topic-triggered translation model for statistical machine translation," *Chinese Journal of Electronics*, vol. 26, no. 1, pp. 65–72, 2017.