Research Article

Penghua Dang*

The extraction method used for English–Chinese machine translation corpus based on bilingual sentence pair coverage

https://doi.org/10.1515/comp-2023-0107
received May 11, 2023; accepted November 11, 2023

Abstract: To improve the effect of corpus extraction in bilingual English–Chinese machine translation (ECMT), this article combines a machine learning algorithm with the ECMT corpus extraction method based on bilingual sentence pair coverage. Aiming at the propagation characteristics of medium- and short-range speeches, and employing multi-band speech waveforms in the troposphere, the actual requirements of remote machine translation are combined with the free-space, 1546, and the dual-path models, respectively, to analyze three typical air-to-ground wave propagation models. A simulation is conducted for both predictions and comparisons. In addition, a radio wave propagation model within the line of sight in the troposphere is established. The results suggest that the improved dual-path model is consistent with the propagation characteristics of radio waves in the troposphere. Moreover, the experiments show that the proposed ECMT corpus extraction method based on bilingual sentence pair coverage can play an important role in more accurate English–Chinese translations.

Keywords: bilingual sentence pair, coverage, English–Chinese, machine translation, corpus extraction

1 Introduction

Preprocessing is a relatively complex task for domain-specific translations. For example, the translation of patent documents generally requires special domain-specific translations, such as identification and translation of patent numbers, mathematical expressions, technical terminologies, and other literal translation components. Although preprocessing may seem simple, the derivation of useful and practical insights requires laborious phases. However, the output quality of the preprocessing will also have a great impact on translation results.

The matching-based method is called the dictionary-based word segmentation approach, which matches a Chinese text with word segmentation within the entries of a large enough machine dictionary according to a certain strategy. The main matching methods are called the maximum matching method, reverse maximum matching method, word-by-word traversal method, segmentation mark method, and optimal matching method [1]. On the other hand, obvious advantages of string matching-based methods exist, which are defined as simple and easy to implement. However, they also have shortcomings, which are delineated as intersection ambiguity and combination ambiguity since even the definition of a word is not uniform. Besides, there is no standard word set and a lack of self-learning intelligence.

On the other hand, words are stable combinations, and the more frequently adjacent words appear together in the context of statistical-based methods. The more likely they become, the greater the chance to form a word. Therefore, the probability or frequency of simultaneous occurrence of adjacent words better reflects the credibility of the word. The main statistical models applied are called the meta-grammar model, hidden Markov model, maximum entropy model, and conditional random field [2]. In more specific applications, statistical-based methods are often combined with dictionary-based methods.

The advantages of the dictionary-based method for word segmentation are described as fast and efficient, and statistical methods are combined with context to identify new words and better eliminate ambiguity, to achieve a better word segmentation effect. Another method based on better understanding is called the word segmentation method which employs artificial intelligence. The motivation of the method is to analyze the syntax and semantics concurrently as word segmentation to deal with possible ambiguity. However, the system of word segmentation has been in its initial phase [3].

The initial preprocessing is the formalization and standardization of the corpus. In addition, the content of sentence

* **Corresponding author: Penghua Dang,** School of Literature and History, Longdong University, Qingyang, Gansu, 745000, China, e-mail: dangphua@163.com

pairs in the bilingual parallel corpus is generally uneven. There may be a large number of translation errors and misaligned sentence pairs, so the translation quality becomes poor, or sentence components are missing. The existence of such a large number of sentence pairs will seriously affect the training stage of machine translations too. Therefore, such sentences need to be filtered in the preprocessing [4]. There are sentence pairs with too simple and complex syntax in the bilingual parallel corpus.

Due to their particularity, these sentence pairs will become noisy data and affect the training of the translation model. For example, the word alignment tool does not handle long sentences very well, and the compound sentences will affect the quality of word alignment to a certain extent, resulting in a decline in translation quality. Therefore, the bilingual parallel corpus needs to be optimized based on the content, and wrong sentence pairs, and finally long and short sentence pairs are filtered out [5]. Setting the range of sentence lengths to filter long and short sentence pairs available in the bilingual parallel corpus is conducted by using the ratio of Chinese and English sentence lengths. So, large differences in length between Chinese and English sentence pairs are filtered out. The range of the length ratio and the range of the sentence length used for filtering are artificially specified quantities to measure according to the statistical information on the corpus [6].

Time, specific numbers, person, place, and institution names also could exist in the bilingual parallel corpus or test set, and the identified entities are expressed in the form of a specific variable. Therefore, the introduction of these specific variables is mainly treated as a kind of special words or phrases having diversity and time-varying characteristics, that is, unregistered words or new words account for a large proportion of the corpus [7]. No such a large corpus can cover them. The number of appearances of such entries in the corpus dictionary is finally increased, which results in spending more time and space when both storage and decoding are conducted. However, the variable representation of the same category and the conversion of specific entries into abstract symbols can not only greatly reduce the vocabulary size and the number of phrases, but can also improve the efficiency of the system [8]. In a translation platform, although the coverage is not as good as the proper name recognition system, the library composed of named entities to realize the recognition and translation of proper names can ensure extremely higher accuracy. Thus, the performance of a system translation would be better. The named reality library implemented in the article is collected through the network [9].

Statistical machine translation methodology has been rapidly developed and applied with the fast development, popularization, and application of computer networks and communication technologies. Moreover, maturing bilingual resource acquisition, processing, and application technology has brought better outcomes. Therefore, even though it has led to more international exchanges, the amount of information has increased rapidly, resulting in language barriers and more simultaneous transactions required correctly. The realization of mutual translation between texts in different languages through computers has become one of the widely researched subjects. Thus, machine translation appears to be an effective solution to this problem as a powerful means [10,11].

The processing and optimization of bilingual corpora needs comprehensive efforts. With the continuous increase in the scale of the corpus, the continuous enrichment of computerized translations, and the continuous isomerization of corpus sources, it is no longer possible to process the corpus manually. How to design a more effective algorithm not being affected by different types of noise in the corpus has increasingly become a challenge [12]. The corpus contains several language-related phenomena, knowledge, and complications. Besides, how to design algorithms to extract deep useful information and insights from the corpus could help users reach better-translated texts. When natural language processing using statistical methods is under consideration, the acquisition, processing, optimization, and application of corpus are undoubtedly the “cornerstone” of the development [13]. Therefore, the effectiveness of the corpus is substantial. The same is true for bilingual corpus too. Due to the importance of translation resources such as corpora for machine translation, in addition to studying how to build, optimize, and implement algorithms, researchers have focused on how to obtain and effectively use bilingual corpus resources [14].

The processing based on the content of the corpus is a relatively small task, and it does not fully play a role in the more effective use of the corpus in the subsequent steps. The content of sentence pairs in the bilingual parallel corpus is generically uneven, so the translation quality of sentence pairs is difficult to guarantee, and there may be many mistakenly translated and misaligned sentence pairs [15]. There are sentence pairs with overly simple and complex syntax in the bilingual parallel corpus. Due to their particularity, these sentence pairs will be assessed as noisy data and affect the training of the translation model. To this end, the content of bilingual parallel corpus needs to be optimized, so automatic methods are used to reasonably evaluate the quality of sentence pairs, and then sentence pairs with poor quality are filtered out. Hence, the translation quality of the system to a certain extent improved [16].

Many of the current bilingual corpora are obtained through automatic methods. Regardless of the parallel or bilingual web pages with mixed sources, automatic methods will be used in the crawling, parsing, sentence alignment, and linking of web pages with other entire parallel corpora, which will inevitably introduce errors and noises, such as sentence segmentation and alignment errors. Generally, it is difficult to guarantee the symmetry and a higher translation quality of the bilingual content of the original webpage. So, it is necessary to evaluate the quality of each sentence pair in the initially obtained bilingual corpus [17]. An effective method needs to be found to filter the sentence pairs with poor qualities in the original corpus. Also, a reasonable scoring mechanism is needed to distinguish the quality of sentence pairs, and these scores are used to rank. While high-quality sentence pairs are ranked first, poor-quality sentence pairs are ranked later [18].

The simplest method uses a bilingual dictionary to evaluate the quality of the sentence pairs in the corpus from the perspective of the mutual translation of the bilingual sentence pairs called loyalty. The evaluation of the quality of sentence pairs will be helpful when sufficient dictionary coverage is available. The results of dictionary-based sentence pair evaluation methods are generally affected by two factors. The first one is the size and coverage of the vocabulary; the other is whether the domain information of the vocabulary is obvious. When the first indicator is not ideal, the result of the evaluation is also unreasonable. Generally, when using this indicator to evaluate the quality of sentence pair translation, the coverage of the source words of the vocabulary to the source words of the candidate corpus is counted, and this method is used only when the coverage is not very narrow. In addition, it is also a good method to select the corpus that is more relevant in a certain field when the mixed corpus is available and has a relatively sufficient dictionary for field translation.

This article studies the English–Chinese machine translation (ECMT) corpus extraction method based on bilingual sentence pair coverage using a machine learning algorithm, builds an intelligent model, and improves the quality of English–Chinese translation.

The rest of the article is outlined as follows: Section 2 presents the preliminary. Section 3 is allocated to the proposed method. Section 4 concludes the research.

2 Modeling and simulation of the waveform propagation of a typical speech

A typical speech as a wave propagation simulated and modeled in the literature is investigated by using free space propagation and dual path models.

2.1 Free space propagation model

If the wave source of the speech waveform is uniformly radiated outward as a spherical wave in the free space and the radiation power is P_Σ, then the energy density at the distance d from the speech sender is expressed by

$$S = \frac{P_\Sigma}{4\pi d^2}. \tag{1}$$

At the place far from the speech originator d, the radiated electromagnetic wave can be considered a uniform plane wave. The ratio of the electric field strength to the magnetic field strength is denoted by

$$E_0/H_0 = 120\pi,$$

where 120π denotes the air resistance and its unit is denoted by Ω, and the electric and the magnetic fields are in the same phase. The average power passing through a unit area is represented by

$$S = E_0 \cdot H_0 = \frac{E_0^2}{120\pi}. \tag{2}$$

In equation (2), E_0 and H_0 are measured values. From equations (1)–(3) is derived.

$$E_0 = \frac{\sqrt{30P_\Sigma}}{d}. \tag{3}$$

In practical applications, equation (3)'s decibel value is mostly used to obtain equation (4).

$$E_0\ (\mathrm{dB}(\mu\mathrm{V/m})) = 104.76 + 10\lg P_\Sigma - 20\lg d. \tag{4}$$

In equation (4), P_Σ denotes the radiated power, the unit is kW, d is the propagation distance, and the unit is km.

Figure 1(a) and (b) is the simulation results of field strength and propagation loss of the free space model with a frequency of 100 MHz, respectively. The horizontal axis is the propagation distance d of the radio wave by using logarithmic coordinates. The vertical axis is the field strength and propagation loss at the receiving point.

The speech height correction of a receiver/mobile station can be calculated. The terrain clearance angle is the elevation angle of the line of sight at the receiver of the speech sender. It is generally denoted as θ_{tca}, and θ_{tca} is generally contained between +0.55° and +40°.

$$\tan\theta_{\mathrm{tca}} = \frac{h - A}{d}. \tag{5}$$

Equation (6) defines L_{angle} as follows:

$$L_{\mathrm{angle}}(\mathrm{dB}) = J(v') - J(v). \tag{6}$$

The final field strength is obtained by adding the correction term to the basic field strength value obtained after

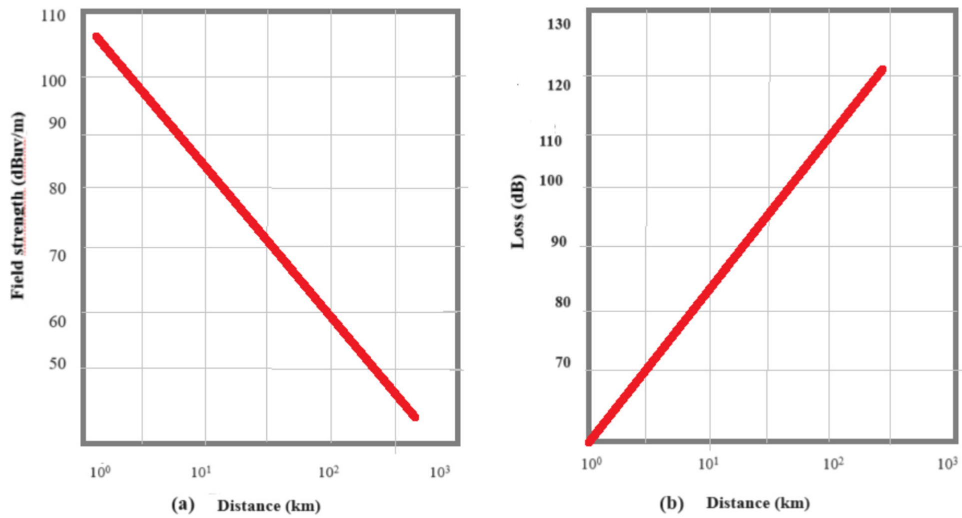

Figure 1: Free space simulations: (a) simulation results of a free space field strength, and (b) simulation results of a free space loss (f = 100 MHz).

interpolation and extrapolation. Equation (7) defines it as follows:

$$E = E_{\text{base}} + L_{\text{an}} + L_{\text{angle}}\ \text{dB}(\mu\text{V/m}). \tag{7}$$

However, it is important to note that for land paths the field strength must not exceed the maximum value $E_{\max}$ defined by

$$E_{\max} = E_{\text{fs}}\text{dB}(\mu\text{V/m}). \tag{8}$$

In equation (8), E_{fs} denotes the free-space field strength of 1 kW effective radiated power, which is given by

$$E_{\text{fs}} = 106.9 - 20\ \text{lg}(d)\text{dB}(\mu\text{V/m}). \tag{9}$$

The transmit power discussed above is all 1 kW, and the field strength expression when the transmit power is not 1 kW is defined by

$$E = E_{1\,\text{kw}} + 10\log(P)\text{dB}(\mu\text{V/m}). \tag{10}$$

The equivalent basic transmission loss at a given field strength is given by

$$L_{\text{b}} = 139.3 - E + 20\ \text{lg}\, f - |D_1 - D_2\ \ (\text{dB}). \tag{11}$$

Based on equation (11), the propagation loss of the 1546 model can be compared to that of other models.

Figure 2 shows the simulation results of the 1546 model corresponding to the heights of different transmitted speech senders at 100 MHz. The horizontal axis denotes the radio wave propagation distance d, and the endpoints of the curves are different because the transmission within the line-of-sight range is considered. The higher the line-of-sight distance is, the higher the line-of-sight distance would be for the sender and receiver, respectively.

Figure 3 depicts the field strength and propagation loss when the height of the transmitted speech sender is 3 km, and the frequencies are, respectively, 50, 100, 200, and 600 MHz. Also, Figure 3 provides an effective basis for the finally established propagation model within the tropospheric line of sight.

2.2 Dual-path model

The double-path propagation model of radio waves on smooth ground is a relatively basic model for inland mobile communication systems. Figure 4 shows a reflection on a smooth surface.

Assumed that the height of the transmitting speech sender A is h_1, the height of the receiving speech sender B is h_2, and I is the lengths of both O_A and O_B. d_1 and d_2 are the direct wave path and the reflected wave path, respectively, and the projection angle between the incident wave and the ground is y. Δr is the distance between the two paths, which can be expressed as:

$$\begin{aligned}\Delta r = d_2 - d_1 &= \sqrt{(h_1 + h_2)^2 + l^2} - \sqrt{(h_1 - h_2)^2 + l^2}\\ &\approx \frac{2h_1h_2}{l}.\end{aligned} \tag{12}$$

The field strength at the receiving point B should be the superposition of the direct wave and the ground-reflected wave. If the amplitude of the field strength produced at receiving point B along the path d_1 is E_1 and the field strength produced at receiving point B along path d_2 is E_2, then the total field strength at B is expressed by

$$\vec{E} = \vec{E}_1 + \vec{E}_2 = E_1(1| + |R|e^{-j(\beta\Delta r + \varphi)}). \tag{13}$$

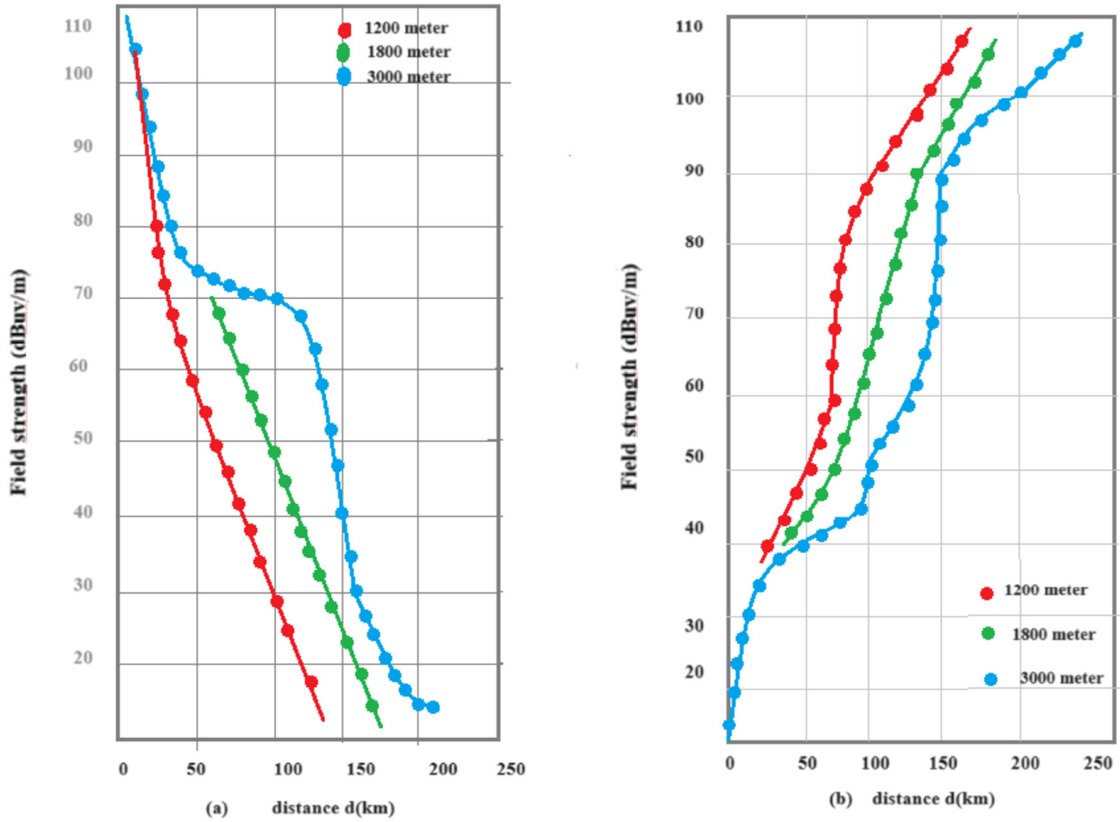

Figure 2: The simulation results of the heights of the different transmitted speech senders: (a) The field strength corresponding to the height of different transmitters of the 1546 model (100 MHz). (b) The loss corresponding to the height of the 1546 model for different transmitted speech senders (100 MHz).

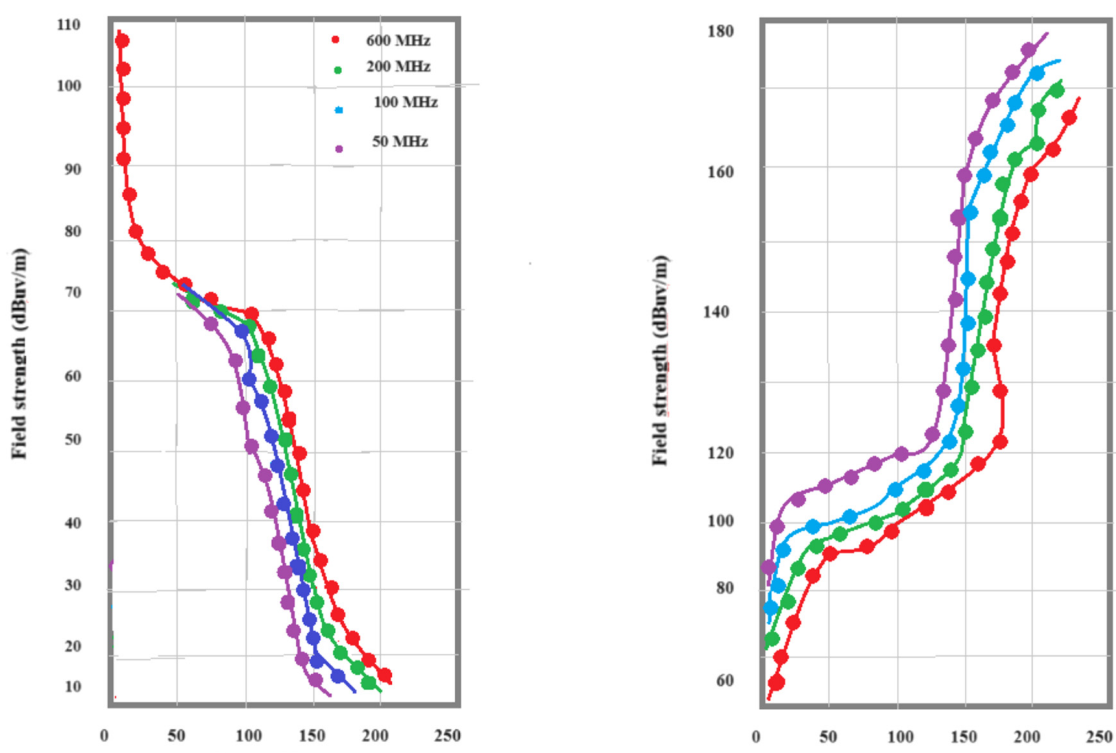

Figure 3: The simulation of the 1546 models with different frequencies: (a) The propagation field strength of the 1546 model at different frequencies when h_1 = 3 km. (b) The propagation loss of the 1546 model at different frequencies when h_1 = 3 km.

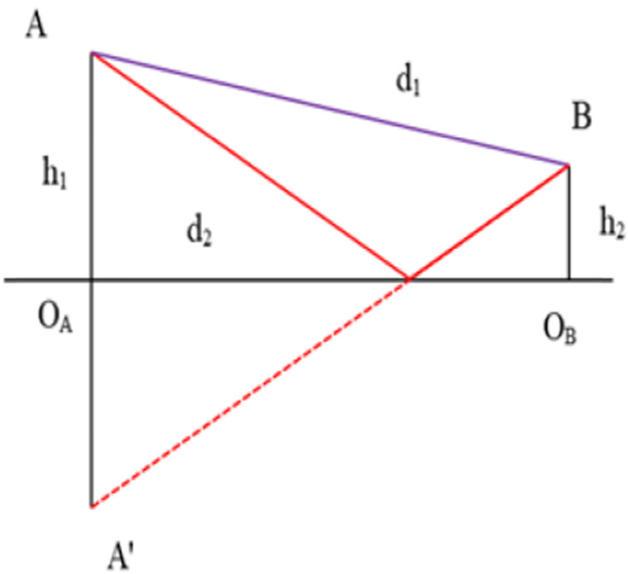

Figure 4: A reflection on a smooth surface.

For horizontally polarized waves, equation (14) is employed.

$$R_H = \frac{\sin\gamma - \sqrt{(\varepsilon_r - j60\lambda\sigma) - \cos^2\gamma}}{\sin\gamma + \sqrt{(\varepsilon_r - j60\lambda\sigma) - \cos^2\gamma}}. \quad (14)$$

In equation (14), ε_r denotes the relative permittivity of the earth, and σ denotes the electrical conductivity of the earth.

For horizontally polarized waves, the actual ground reflection is closer to the ideal conductive ground, especially in regions with longer wavelengths or smaller projection angles. Therefore, when the influence of ground reflection is estimated, the actual ground can be roughly equivalent to an ideal conductive ground.

For vertically polarized waves, equation (15) is employed.

$$Rv = \frac{(\varepsilon_r - j60\lambda\sigma)\sin\gamma - \sqrt{(\varepsilon_r - j60\lambda\sigma) - \cos^2\gamma}}{(\varepsilon_r - j60\lambda\sigma)\sin\gamma + \sqrt{(\varepsilon_r - j60\lambda\sigma) - \cos^2\gamma}}. \quad (15)$$

In equation (15), ε_r denotes the relative permittivity, which is 4 in dry soil and 80 in seawater, and σ denotes the electrical conductivity.

The final ground reflection coefficient is denoted by

$$|R| = \sqrt{R_H^2 + R_V^2}. \quad (16)$$

When γ is small, the resultant field can be simplified as follows:

$$|E| = |E_1 + E_2| = 2E_1\left|\sin\left(\frac{2\pi h_1 h_2}{\lambda l}\right)\right|. \quad (17)$$

When the communication distance is large, equation (17) denotes the resultant field that cannot be directly applied, and the effect of the earth's curvature must be considered.

There are two effects of Earth curvature on the dual-path model. First, when the field strength of the receiving point is calculated by the superposition method of the direct wave and the reflected wave, the path difference between the direct wave and the reflected wave on the spherical ground is different regarding the plane ground, so the heights h_1 and h_2 of the speech sender must be properly corrected. Second, the sphere has a diffusion effect on the reflection of the radio wave, so the change in the electric field intensity caused by the diffusion effect must be considered.

Figure 5 shows that when the propagation distance is large, the influence of the curvature of the earth on the propagation of radio waves must be considered. The heights h_1 and h_2 of the speech sender should be properly corrected; that is, the equivalent speech sender heights h_1' and h_2' should be used instead. The tangent plane is made through the reflection point C, and then, the equivalent height of the sender at this time should be the vertical height h_1', and h_2' that is calculated from the tangent plane. Because the angles between h_1 and h_1' and h_2 and h_2' are very small, equation (18) is used to represent approximations as follows:

$$h_1' \approx h_1 - \Delta h_1, \quad h_2' \approx h_2 - \Delta h_2. \quad (18)$$

According to the derivation method of the line of sight distance, equations (19) and (20) can be obtained as follows:

$$\Delta h_1 \approx \frac{d_1^2}{2R_0}, \quad (19)$$

$$\Delta h_2 \approx \frac{d_2^2}{2R_0}. \quad (20)$$

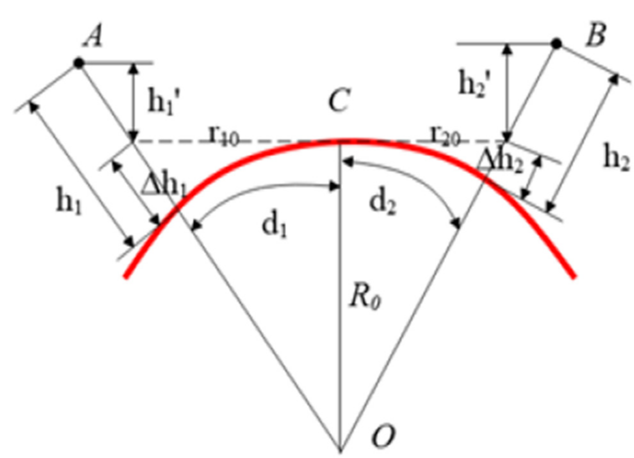

Figure 5: Equivalent height of the speech sender.

Then, the equivalent height of the sender on the spherical ground is defined by

$$h_1' \approx h_1 - \frac{d_1^2}{2R_0}, \tag{21}$$

$$h_2' \approx h_2 - \frac{d_2^2}{2R_0}. \tag{22}$$

In equation (22), $R_0 = 6{,}370\,\text{km}$ is the radius of the earth, and h_1' and h_2' are the corrected heights of the senders of the transmitted and received speech.

To calculate a line-of-sight propagation, it is one of the corrections under spherical ground conditions if the actual height of the speech sender is replaced with an equivalent height.

The field strength of the reflected wave after the radio wave is reflected on the spherical surface is smaller than that on the plane ground. Due to the existence of diffusion, the reflection coefficient of spherical ground is smaller than that of plane ground of the same geology. The physical quantity that describes the degree of diffusion is called the diffusion factor. The diffusion factor of spherical ground is defined as follows:

$$D_f = \frac{\text{Spherical ground reflection reflected wave field strength}}{\text{Plane ground reflection reflected wave field strength}}. \tag{23}$$

Equation (24) presents it.

$$D_f = \frac{1}{\sqrt{1 + \frac{2d_2 d_1^2}{KR_0 d h_1'}}}. \tag{24}$$

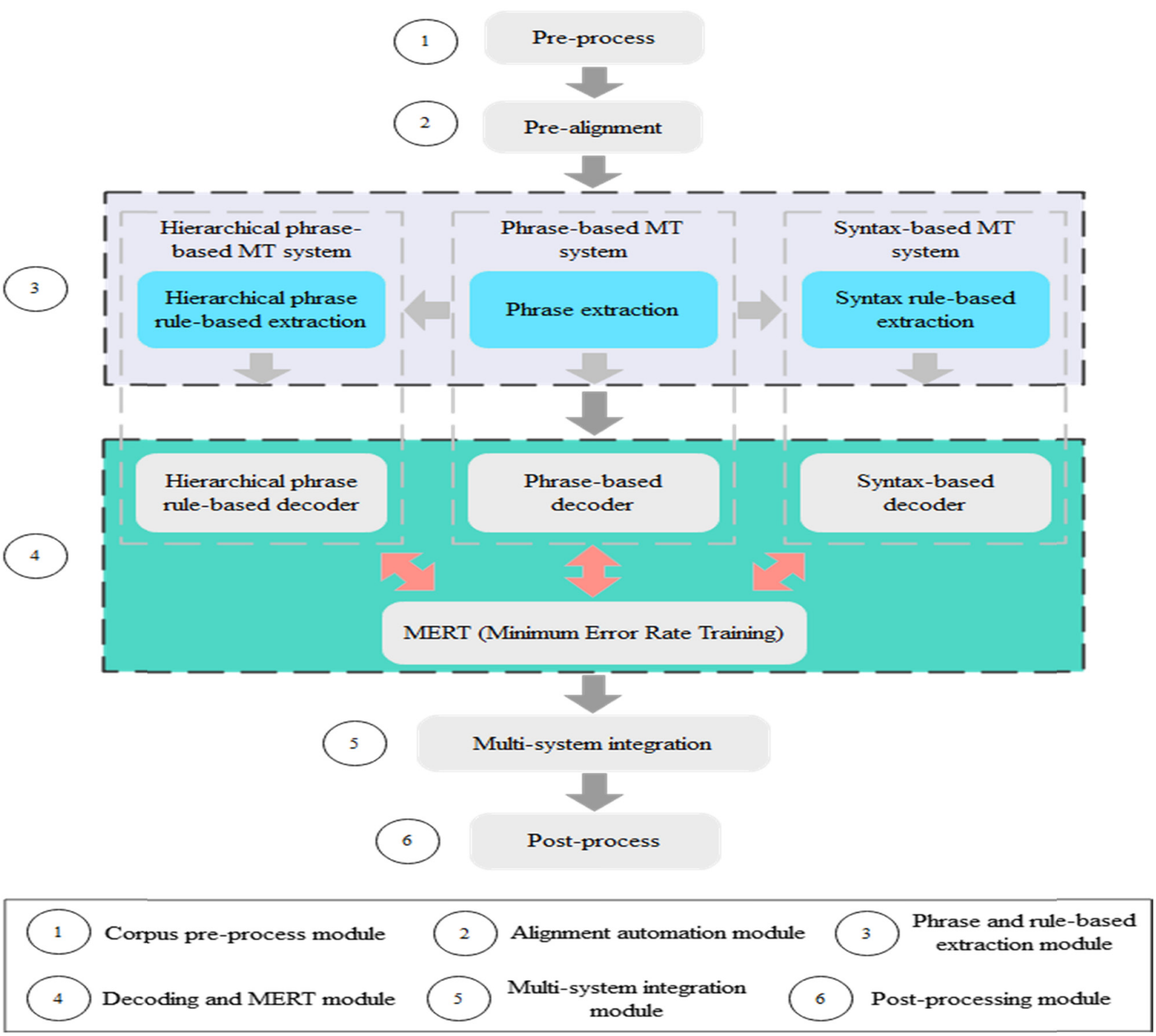

Figure 6: The framework of the ECMT platform.

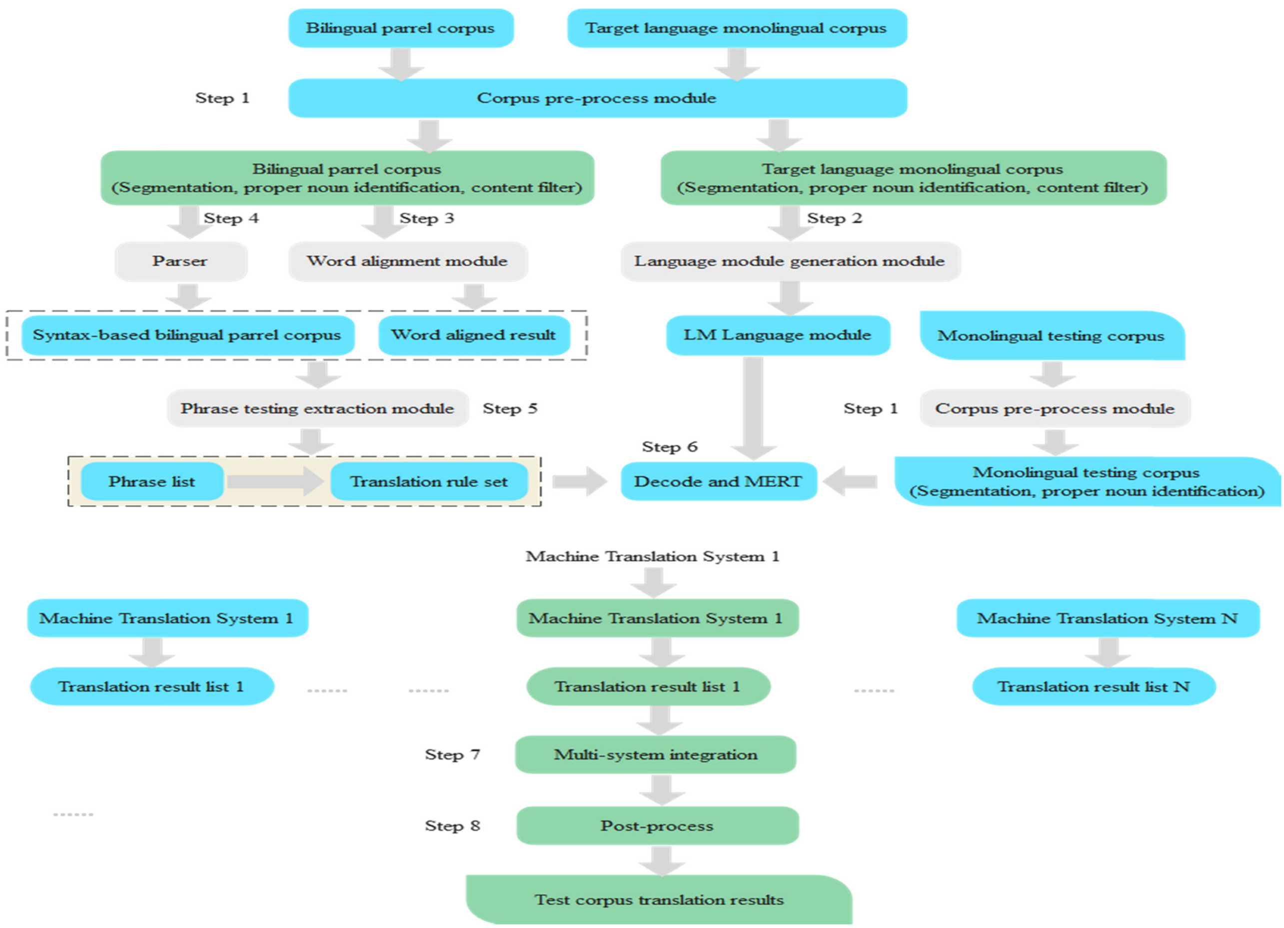

Figure 7: The translation flow chart of the ECMT platform.

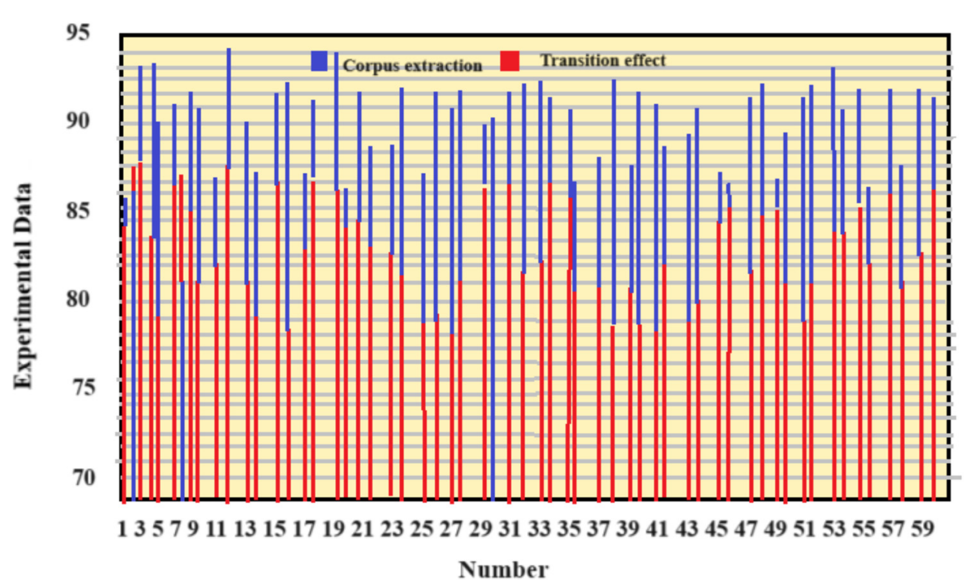

Figure 8: The verification of the extraction method's effect on the ECMT corpus based on bilingual sentence pair coverage.

After introducing the diffusion factor, it is necessary to replace the reflection coefficient R in the calculation of line-of-sight propagation D_f. R completes another correction under the condition of the curved surface.

The final field strength E is defined by

$$E = \frac{173\sqrt{P_\Sigma D_1 D_2}\sqrt{1 + D_f^2\,|R|^2 + 2D_f|R|\cos\left(\frac{2\pi\Delta r}{\lambda} + \phi\right)}}{l}. \tag{25}$$

3 Extraction method of the ECMT corpus based on bilingual sentence pair coverage

This article constructs a multi-engine statistical machine translation platform, which is mainly composed of a phrase-based translation system, a hierarchical phrase-based translation system, and a syntax-based translation system.

A platform based on both automation and modularization is designed. Automation mainly refers to reducing manual intervention as much as possible when preprocessing to model training, translation generation based on test sets, and automatic parameter adjustment are conducted by computers. The framework of the ECMT platform is shown in Figure 6.

The basic translation process of a statistical machine translation platform consists of the following steps, which include corpus preprocessing, language model generation, word alignment, syntactic analysis, phrase and rule extraction, decoding, and post-processing. Figure 7 shows the basic flow chart of a bilingual translation.

The effect of the proposed extraction method of the ECMT corpus based on bilingual sentence pair coverage is verified. The corpus extraction and translation effects are counted. Figure 8 shows the outcomes.

Figure 8 depicts that the proposed extraction method of the ECMT corpus based on bilingual sentence pair coverage can play an important role in bilingual English–Chinese translation tasks.

4 Conclusion

Corpus preprocessing is an important fundamental phase in statistical machine translation, and the quality of preprocessing is very related to the cost of the training used for machine translation and the translation effect. In the preprocessing module, operations such as Chinese word segmentation, full-width and half-width conversion, proper name recognition, and translation are generally performed on a corpus to generate the qualified corpus that is used for machine translation systems. In addition, to reduce the size of the vocabulary of the corpus used for training, speed up the training stage, and improve the quality of word alignment to a certain extent, some specific types of entities identified will be generalized during preprocessing.

The manuscript combines machine learning algorithms with the proposed extraction method called the ECMT corpus based on bilingual sentence pair coverage to construct an intelligent model. The experimental results show that the proposed extraction ECMT corpus based on the bilingual sentence pair coverage method can play an important role in bilingual English–Chinese translation.

The future direction of the research will be based on other advanced machine learning methods to better obtain results in bilingual machine translations and increase the efficiency of bilingual English–Chinese translation tasks.

Funding information: This study did not receive any funding.

Author contributions: The full research is conducted by the author.

Conflict of interest: The authors declare that they have no conflict of interest.

Ethical approval: No consent is required.

Informed consent: Not necessary.

Data availability statement: Data is available upon request.

References

[1] Z. Sun, M. Anbarasan, and D. J. C. I. K. Praveen, "Design of an online intelligent English teaching platform based on artificial intelligence techniques," *Computat. Intell.*, vol. 37, no. 3, pp. 1166–1180, 2021.
[2] H. Kim, "The efficacy of Zoom technology as an educational tool for English reading comprehension achievement in EFL classroom," *Int. J. Adv. Cult. Technol.*, vol. 8, no. 3, pp. 198–205, 2020.
[3] A. E. P. Atmojo and A. Nugroho, "EFL classes must go online! Teaching activities and challenges during the COVID-19 pandemic in Indonesia," *Regist. J.*, vol. 13, no. 1, pp. 49–76, 2021.
[4] A. Yuliansyah and M. Ayu, "The implementation of project-based assignments in online learning during COVID-19," *J. Engl. Lang. Teach. Learn.*, vol. 2, no. 1, pp. 32–38, 2021.
[5] S. Liu and J. Wang, "Ice and snow talent training based on construction and analysis of artificial intelligence education informatization teaching model," *J. Intell. Fuzzy Syst.*, vol. 40, no. 2, pp. 3421–3431, 2021.

[6] E. Apriani and J. Hidayah, “The ICT used by the English lecturers for non-English study program students at STAIN Curup,” *Vision J. Lang. Foreign Lang. Learn.*, vol. 8, no. 1, pp. 26–37, 2021.
[7] R. Godwin-Jones, “Second language writing online: An update,” *Lang. Learn. Technol.*, vol. 22, no. 1, pp. 1–15, 2018.
[8] P. Pérez-Paredes, G. C. Ordoñana, and J. P. Aguado, “Language teachers’ perceptions on the use of OER language processing technologies in MALL,” *Comput. Assist. Lang. Learn.*, vol. 31, no. 5–6, pp. 522–545, 2019.
[9] Y. Bin and D. Mandal, “English teaching practice based on artificial intelligence technology,” *J. Intell. Fuzzy Syst.*, vol. 37, no. 3, pp. 3381–3391, 2019.
[10] R. W. Todd, “Teachers’ perceptions of the shift from the classroom to online teaching,” *Int. J. TESOL Stud.*, vol. 2, no. 2, pp. 4–16, 2020.
[11] M. A. Al-Ghamdi, “Arabic learners’ preferences for Instagram English lessons,” *Engl. Lang. Teach.*, vol. 11, no. 8, pp. 103–110, 2018.
[12] Z. Xu and Y. Shi, “Application of constructivist theory in flipped classroom college English teaching as a case study,” *Theory Pract. Lang. Stud.*, vol. 8, no. 7, pp. 880–887, 2018.
[13] G. A. Toto and P. Limone, “Motivation, stress, and impact of online teaching on Italian teachers during COVID-19,” *Computers*, vol. 10, no. 6, pp. 75–94, 2021.
[14] C. Meng-Vue, L. Dan, and W. Jun, “A study of college English culture intelligence-aided teaching system and teaching pattern,” *Engl. Lang. Teach.*, vol. 13, no. 3, pp. 77–83, 2020.
[15] Y. Rinantanti, S. Z. Bin-Tahir, and A. Suriaman, “The impact of EFL senior high school teachers’ performance in Papua, Indonesia toward the students’ English learning achievement,” *Asian EFL J.*, vol. 23, no. 3.3, pp. 431–447.
[16] R. Shadiev and M. Yang, “Review of studies on technology-enhanced language learning and teaching,” *Sustainability*, vol. 12, no. 2, pp. 524–532, 2020.
[17] M. A. S. Khasawneh, “An electronic Training Program on Developing Written Expression Skills among a Sample of foreign language learners EFL who are at risk for Learning disabilities during the emerging COVID-19,” *Acad. Soc. Sci. J.*, vol. 7, no. 10, pp. 1974–1982, 2021.
[18] N. Hollenstein, J. Rotsztejn, M. Troendle, A. Pedroni, C. Zhang, and N. Langer, “ZuCo, A simultaneous EEG and eye-tracking resource for natural sentence reading,” *Sci. Data*, vol. 5, no. 1, pp. 1–13, 2018.