ISSN:2790-167X DOI: 10.56028/aehssr.3.1.284

Machine Translation Post-Editing Training Room Construction

Xiao Xiao ,Hui Sun* ,Yanhua Wang

School of Foreign Languages, Shanghai Technical Institute of Electronics & Information, Shanghai, China

xiaoxiao@stiei.edu.cn, halleysunny@163.com, wangyanhua@stiei.edu.cn

*Corresponding Author

Abstract. This paper, based on the transformation and development of applied English majors in vocational undergraduate education, is intended to construct machine translation post-editing training rooms by analyzing the language industry development status, foreign languages majors' competence adaptability for enterprises requirements, and the connotation of machine translation post-editing training with a view to provide suggestions and references for the construction of applied English practical courses.

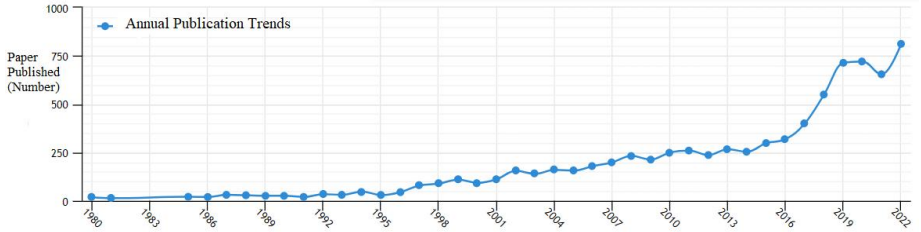
Keywords: vocational undergraduate education; machine translation post-editing; construction of training rooms

1. Introduction

In early 2021, the introduction of the Administrative Measures on the Establishment of Undergraduate Vocational Education Majors (Trial) (hereinafter referred to as the "Measures") marked the official launch of the Ministry of Education on the establishment of undergraduate majors in vocational colleges and universities and the formal establishment of the national system for vocational undergraduate education majors. It also marks that after years of demonstration, vocational undergraduate education has finally shifted from research to comprehensive practice. Under the guidance of the Measures, the Ministry of Education has adjusted and set professional standards for Applied English Majors in vocational undergraduate education. The professional direction has been transformed. Machine translation and post-editing courses have been integrated into professional core courses and internship training ones, filling the gap in the curriculum system where traditional language majors have long been disconnected from language technology. This change is bound to usher in new opportunities and challenges for the construction of language training rooms.

The construction of a post-editing training room targets building and accumulating bilingual parallel corpus with localized characteristics according to the school's positioning and characteristics. Meanwhile, the team organizes large-scale machine translation post-editing competitions to carry out social services and training such as post-editing curriculum construction, post-editing competition and other social services and training for other peers with a purpose to expand the social influence and economic benefits of our college within three years progressively.

Fig 1 Research trend of machine translation in recent 40 years



It can be easily recognized that machine translation has shown an overall upward trend in the past 40 years according to Fig 1, and machine translation entered a boom period in 2006 when the

ISSN:2790-167X DOI: 10.56028/aehssr.3.1.284
key word “machine translation” is input in the CNKI. Especially since 2016, relevant researches have shown an explosive growth, and machine translation has been an important direction for the transformation of professional undergraduate English majors.

2. Transformation and development

In February 2022, extensive comments on professional profiles for secondary vocational education, higher vocational colleges and vocational undergraduates were solicited by the Ministry of Education, focusing on applied English majors offered by both vocational colleges and undergraduates. In comparison, it could be found that vocational undergraduates have systematically sorted out and put forward different requirements in terms of career orientation, training target positioning, professional ability requirements, professional courses and internship training, vocational certificates and continuing courses. In terms of career orientation, the characteristic position of “translator” language talent is retained in the vocational undergraduates, making “administrative affairs processing staff” focus on the business field of “business professionals, sales staff, business consulting service personnel”. The post group adjustment is more reasonable, moving from loose to typical positions clearly, i.e., vocational education undergraduates are oriented towards language services, foreign-related trade services and management, reflecting the feature of “occupation of vocational education”. In terms of training target positioning, the term “better” was adjusted to that of being “solid” and “good” was adjusted to “stronger”, and requirements such as “on-site interpretation and data translation”, “machine translation post- editing” and “high-level technical skills” were put forward, etc., and more emphasis was placed on professional ability training requirements such as “on-site interpretation” and “machine translation post-editing”. High-level in vocational undergraduate education can be seen easily. Therefore, in terms of professional courses and internship training, according to the specific connotation of career orientation and training target positioning, the basic vocational undergraduate course proposed “English Translation Theory”, and added “Machine Translation Post- Editing” to the core courses. The practical project training of “Machine Translation Post-Editing” is increased marginally in the internship training session, highlighting the development of translation skills, building a platform for students to contact real projects of enterprises during school years, and reinforcing the positioning of technical skills in English language learning. Vocational certificates and continuing courses have a clearly direction, listing the direction for China Accreditation Test for Translators and Interpreters (CATTI) and Professional Masters in Translation and Interpretation. Therefore, from the perspective of the overall setting and adjustment plan for the applied English professional standards in vocational undergraduate education, the direction of translation becomes clearer. Translation, as the main line of professional ability development, has urgent requirements for training translation technical skills. The “Machine Translation Post-Editing” course is designed in the professional core curriculum and practical training course, aiming to adapt to the new development trends of China’s language service industry and the demand for high-level technical and skilled talents in vocational undergraduate education[1].

3. The connotation of machine-based post-translation editing

3.1 Status Quo of the Language Service Industry

Since its inception in the 1930s, machine translation has evolved and updated, experiencing the period of creation, frustration, recovery, development and prosperity, and by 2015, the rise of neural network machine translation (NMT) became a milestone in the history of machine translation. In 2016, Google launched Google Neural Machine Translation (GNMT). Within two years, well-known Internet companies such as Microsoft, Sogou, and NetEase also launched their own neural network machine translation engines, and NMT quickly became a popular language service

ISSN:2790-167X DOI: 10.56028/aehssr.3.1.284
product. Research shows that the output quality of neural network machine translation has improved significantly in comparison with statistical machine translation based on a bilingual parallel corpus training translation model, which has contributed to a significant improvement in translation quality, but there are still many problems inherent in machine translation. GNMT has significantly improved the quality of machine translation translations, but common problems still cannot be avoided such as mistranslation, omissions, contextual inconsistencies, and inconsistent terminology in machine translations, which are more prominent in more detailed vertical domains[2]. Up to now, researches on recognition oriented method on a computationally limited platform have yielded achievements in computer language learning[3][4][5]. Therefore, manual intervention in post-editing is an unavoidable and important link.

3.2 Competence Requirements for Foreign Language Talents

With the integration of information technology and language services, there have also been new changes in the competency requirements of translation talents from large enterprises and translation companies. Table I shows two translator recruitment positions.

Table 1.Translator recruitment requirements

A listed technology company recruits an English translator for translating online texts. The job requirements are as follows: 1. English Majors background, holding TEM-8 or proficiency in English or excellent translation skills; 2. Familiar with translation trends and mainstream CAT tools, familiar with the translation mode of MTPE, and relevant translation experience is preferred; 3. Fond of network literature or comics, having relevant translation experience or other translation experience in vertical domains and the construction of corpus glossary experience is preferred; 4. Good at communication, good pressure resistance and teamwork ability, strong sense of responsibility, priority for managing translation team.
A translation company recruits “Machine Translation + Post- Editing” (MTPE) translators. The specific requirements are as follows: 1. A formal university degree or above; a good foundation of Chinese and English language, with more than 1 year experience in translation; 2. Proficient in using mainstream MT engines and CAT tools for efficient machine translation; 3. Excellent awareness of post- editing, able to apply light PE or heavy PE to the translation as required to meet the requirements of the company and customers in terms of legibility, accuracy, style and format of the translation; 4. Adequate and flexible working hours, university teachers/students/freelance translators are preferred; Please provide convenient, reliable and diverse contact information; 5. Good service awareness and professional ethics, punctual delivery, humility and timely handling of customer feedback, responsible for their own translations.

As can be seen from two recruitment posts above, both translators and specialized MTPE translators have high requirements for familiarity with CAT tools and PE models. It can be seen that the proportion of technical skills in translation work is rapidly increasing, and currently skilled talents have become an urgent need for the development of the language service industry. On the other hand, in addition to the bilingual cultural skills and learning competences, post-editors also need to fully understand the working principles of machine translation, getting familiar with common mistakes in machine translations, having awareness of post-editing, and proficiency in post- editing tools. Therefore, in response to market demand, how to effectively set up the

ISSN:2790-167X DOI: 10.56028/aehssr.3.1.284
“Machine Translation Post-Editing” course for vocational undergraduate education is significantly important[6].

3.3 The Connotation of Machine Translation Post-Editing

Post-editing (PE for short) is the process of modifying and processing the initial output of machine translation (Allen, 2012)[7]. The national standard (GB/T 40036-2021/ISO 18587:2017)[8] defines that post-editing could be realized on the basis of machine translation results and aims to check the accuracy and intelligibility of machine translation, improve text and its readability, and correct errors. The goal of post-editing is recipient or consumer-oriented, that is to say, as long as the recipient or consumer or translation user believes that the quality of the translation can meet their own requirements, the quality judgment standard of post-editing becomes more resilient. Generally, the types of post-editing can be distinguished according to the degree of human intervention in machine translation. There are two main categories of full post-editing and light post-editing. The results of full post-editing should be accurate, understandable, and appropriate in style, with correct use of syntax, grammar and punctuation, and the aim of producing translations that have the same effect as human translation results[9]; while light post-editing will use as much of the original machine translation results as possible to ensure that no information has been added or omitted, any inappropriate content has been modified, and sentence structures have been restructured where the meaning is incorrect or unclear. Vocational undergraduate talents focus on integrating technical skills management and service ability development in professional fields[10]. The authors believe that the post-editing in the new course “Machine Translation Post-Editing” focuses on the proofreading of machine translation, which is closer to light post-editing in the post-editing category, that is, using machine translation as much as possible to correct obvious errors and make the text understood easily, mainly revising mistranslation, cultural differences in content, and reorganizing sentence structure. Light post-editing is the post-editing with the least manual intervention. Therefore, the post editing mentioned in this paper can be understood as light post-editing.

4. Machine translation after editing training room construction

4.1 General Ideas

At present, foreign language laboratories are mostly featured with traditional audio rooms. Later, due to the continuous development of foreign language disciplines and information technology upgrading, simultaneous interpretation rooms, translation laboratories, lecture rooms and other training rooms have been put into construction and application one after another. However, with the advent of the “big data era”, a wave of intelligent online education is spreading and sweeping the world, traditional teaching mode has been transferring to MOOC, cloud classroom and other new teaching modes. In 2020, COVID-19 Pandemic broke out and accelerated the transformation of teaching models to the cloud. Therefore, it is obvious that the traditional translation teaching model cannot meet the needs of modern foreign language teaching media. The software in the existing simultaneous interpretation training room is not enough to support the college's translation services to take modular courses, while computers, the Internet, cloud computing and other technologies can inject new vitality into language and translation teaching, empowering the cultivation of high-quality and highly skilled language talents.

Therefore, the machine translation post-editing training room needs to keep up with the development of emerging technologies. We should establish a correct concept of technological development for students, and cultivate innovative talents with “language + skills” that may meet market needs. Based on modern information education technology, the physical environment of the machine translation post-editing training room can be newly built or upgraded based on the original simultaneous interpretation training room. Special translation software and corpus software need to

ISSN:2790-167X DOI: 10.56028/aehssr.3.1.284
be configured to support mainstream computer-aided translation, machine translation post-editing, translation project management, corpus and other functions, and to support the practical content of translation technology courses and translation cognitive training. Finally, we may establish a positive circular industrial teaching platform, realize integrated translation teaching and practical training, continuously improve the quality of translation teaching, expand translation practice channels, improve the market adaptability and employment advantages of foreign language graduates, empower the development of foreign language disciplines, and prepare students for continuing to study undergraduate translation in the future.

4.2 Construction Goals

The construction of machine translation post-editing training room focuses on the demand for new translation talents in the "artificial intelligence + big data" era, builds an integrated ecosystem of teaching, learning, research and application based on teaching informatization, strives to strengthen the hard-core of professional construction, promote the action of language professional teaching reform, reinforce the talent training, faculty development and construction, reform in-class teaching mode and innovate learning forms outside the classroom, in order to realize the overall goal of cultivating complex, highly qualified, highly skilled and innovative talents with foreign language communication and communication skills, international vision and high humanistic qualities, and improving scientific research and technical service capabilities.

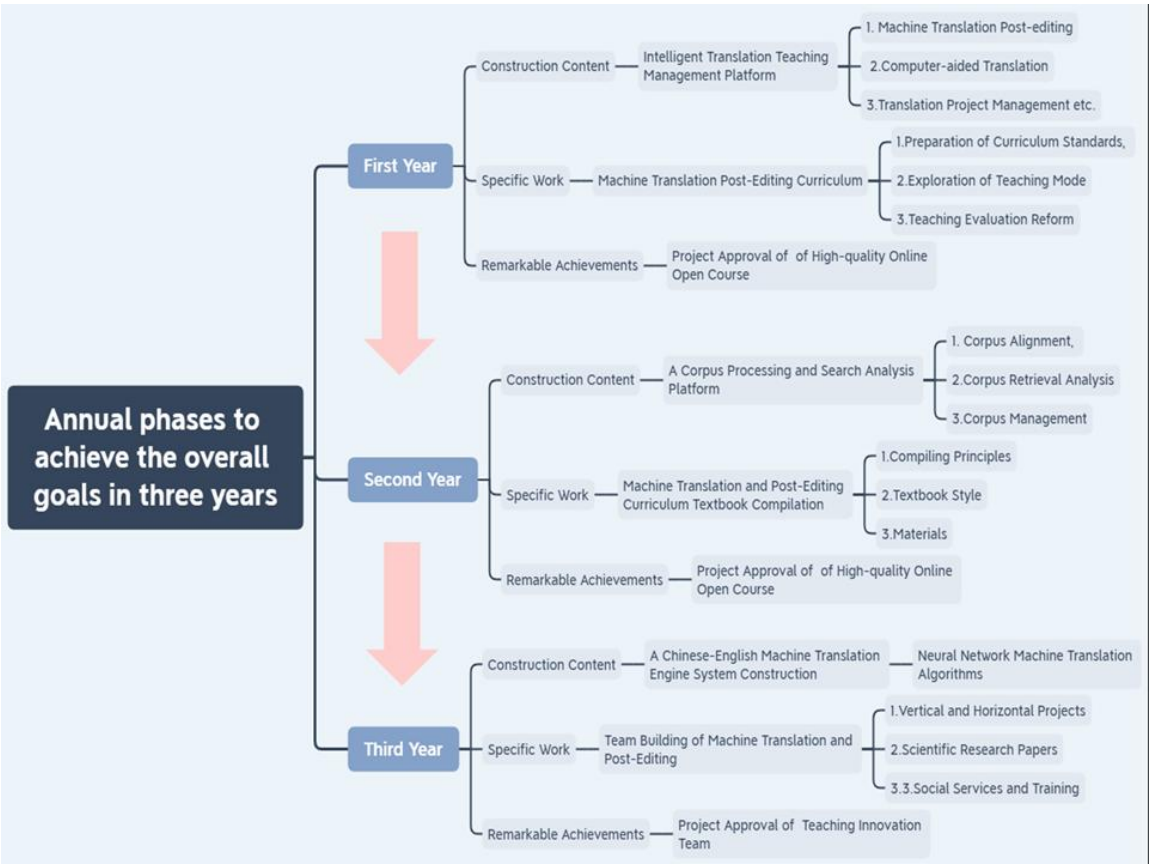
4.3 Construction Path

The construction of a post-editing training room, specifically, includes an intelligent translation teaching management platform, a corpus processing and retrieval management platform, and a machine translation engine system in the professional field[11]. It has teaching and research functions such as machine translation post-editing, computer-aided translation, translation project management, corpus and terminology library, etc., to help teachers and students in translation teaching，research and practice.

It can be built in phases annually to achieve the overall goal in three years according to Fig 2.

Fig 2 Annual phases to achieve the overall goals in three years

ISSN:2790-167X DOI: 10.56028/aehssr.3.1.284



Intelligent translation teaching management and other platforms provide a realistic enterprise translation management and operating environment, making teaching more practical. Student management in teaching, assignment distribution, delivery, correction, feedback and other processes can be completely simulated project team management in the actual translation of enterprises. Project distribution, delivery, review, feedback and other processes can be closely integrated with teaching and practice. Meanwhile, students can familiarize themselves with the actual translation process and operation, as well as the division of labor and responsibilities of different roles in the translation team, with the purpose of learning and applying to the greatest extent. Through the translation memory and terminology management module of the online translation teaching management training platform, teachers and students can easily import and store corpus of different types of files during translation training, and can fully accumulate their own unique achievements such as translation teaching and translation scientific research.

The online translation teaching management training platform has a powerful and stable memory search function, which can quickly search existing corpus and maximize the reuse of language resources. In the translation practice, the platform can also provide teachers and students with a large number of high-quality translation examples to make translation easier. While reducing the difficulty of teaching translation, it also makes students' translation practices more instructive and comparative.

The teaching and research innovation team can use the translation teaching management platform and the corpus search and analysis platform to build and accumulate bilingual parallel corpus with localized characteristics according to the school's positioning and characteristics. At the same time, the team organizes large-scale machine translation post-editing competitions to carry out social services and training such as post-editing curriculum construction, teacher training, student examination and other social services and training for other peers, and further expand the social influence and economic benefits of our college.

ISSN:2790-167X DOI: 10.56028/aehssr.3.1.284

Conclusions

Under the leadership of vocational undergraduate education, professional settings are oriented towards new business formats, new models and new occupations, bringing great prospects to the current needs of the language service industry and the market for machine translation post-editing, and the construction of training rooms needs to be strengthened urgently. “Online + offline”, “in-class + extracurricular”, “theory + practice”, “on-school + off-campus” and virtual simulation teaching, etc., have become the norm for teaching and learning. The construction of machine translation post-translation editing and translation training rooms in higher vocational colleges and universities needs to fully consider comprehensive factors such as post-translation curriculum construction, teaching reform, evaluation mechanism, scientific researches and social service benefits, etc., and actively mobilize to increase spending in the resources of schools, enterprises and school-enterprise double-qualified teachers, students and other relevant parties. Regular feedback and corrections should be given to various issues during the construction and operation process, and gradually a good operation mode will be formed to ensure the sustainability of practical training and practical results.

Acknowledgment

This study is supported by the project of Teaching Guidance Committee for Foreign Language Majors in Vocational Colleges of the Ministry of Education entitled A Study on the Cultivation of Higher Vocational Learners' Metacognitive Ability Based on the Integration of Information Technology and English Curriculum (No. WYJZW-2020-1304, Project Master: Xiao Xiao) and by the project of Shanghai Vocational Education Association entitled Research on Digital Learning Mode in Vocational Education under the Epidemic Normalization Context (No.YB2022012, Project Master: Xiao Xiao)

References

[1] The Ministry of Education of China. Exposure Draft of Professional Profile of Vocational Education, 2022(3).

[2] Zhao Tao. “Current Situation and Problems of Machine Translation Post-Editing,” Foreign Language Teaching, vol.42, issue 4, pp. 100-104, 2021.

[3] B Xue, W Yi, F Jing, et al. “Complex ISAR Target Recognition using Deep Adaptive Learning,” Engineering Applications of Artificial Intelligence, issue 97,pp. 97:1-9,2021.

[4] B Xue, Y He, F Jing, et al. “Dynamic Coarse-to-Fine ISAR Image Blind Denoising using Active Joint Prior Learning,” International Journal of Intelligent Systems, vol 36, issue 8, pp. 4143-4166, 2021.

[5] B Xue, Y He, F Jing, et al. “Robot Target Recognition using Deep Federated Learning,” International Journal of Intelligent Systems, vol 36, issue 12, pp. 7754-7769, 2021.

[6] Liu Chengyou, Feng Liying, Zhao Feng. “Exploration on the Construction of Vocational Undergraduate Course System,” Journal of Hainan Normal University (Natural Science), issue 2, pp. 234-238, 2021.

[7] Allen, J. Post-editing [A]. In Somers, H. (ed.). Computers and Translation: A Translator’s Guide [C]. Shanghai: Shanghai Foreign Language Education Press, pp. 297-317, 2012.

[8] GB/T 40036-2021/ISO18587:2017. Translation Service Post-Editing Requirement of Machine Translation [S].

[9] Feng Quangong, Cui Qiliang. “Post-Editing Research: Focus Analysis and Development trend,” Shanghai Translation, issue 6, pp. 67-74, 2016.

[10] Yang Xinbin. “Thinking and Exploration on Talent Training Mode of Vocational Undergraduate Education,” Research in Higher Education of Engineering, issue 1, pp. 127-133, 2022.

ISSN:2790-167X DOI: 10.56028/aehssr.3.1.284
[11] Zhong Wenming, Shu Chao. “The Competence Structure and Curriculum of Post-Editing -- Based on the Frontier Analysis of Foreign Post-Editing Courses,” Technology Enhanced Foreign Language Education, issue 6, pp. 68-91, 2020.