

Construction of Intelligent Online Tourism English Model Corpus

Jianzhou Cui

Wuxi City College of Vocational Technology, Wuxi, Jiangsu, China

Abstract: In the context of social informatization advancing and competition in the tourism industry intensifying, tourism English service has become a crucial communication tool. Utilizing online corpora, which are built upon language and textual foundations, enables the sharing of linguistic information. As the tourism sector thrives, the importance of tourism English services is increasingly highlighted. To meet the demand for high-caliber talent driven by societal progress, China has implemented ongoing educational reforms, introducing courses like online learning and remote access in domestic colleges and universities. Concurrently, intelligence has emerged as a key research focus in the forthcoming information era. Given this backdrop, the efficient utilization of contemporary information technology to develop an intelligent online tourism English model corpus is a subject of substantial interest. This study employs a method of questionnaire surveys and data analysis to explore the creation of an effective, practical, and compatible intelligent online travel English model corpus based on ESP (English for Specific Purposes) technology. Furthermore, the aim is to better cater to users' personalized resource needs. According to the survey findings, a majority of respondents advocate for the construction of an intelligent online travel English model corpus, recognizing its crucial role in advancing translation work and related fields.

Keywords: ESP; Corpus; Tourism English; Intelligent Model

1. Introduction

There is a noticeable trend in today's society towards advancing and applying information resources. The tourism industry faces increasing market demand and intense competition, prompting numerous domestic

companies to initiate intelligent construction projects to enhance service quality, improve efficiency, and attain a competitive edge. Traditional approaches in the tourism sector are inadequate in meeting diverse consumer needs for high-quality and varied consumption experiences. The creation of an intelligent online English corpus rooted in English for Specific Purposes (ESP) offers an effective solution to these challenges. This article aims to explore the development of an intelligent online corpus based on ESP technology, enabling users to access essential information resources and reference materials more efficiently and accurately. Additionally, it seeks to assist in addressing the existing problems associated with intelligent services.

Considering the current academic standpoint, there exists a plethora of research findings regarding ESP corpus and English corpus construction. For instance, Liu Ping's proposal to conduct educational experiments for acquisition of academic English vocabulary using the ESP corpus has successfully showcased the effectiveness of incorporating a corpus in academic vocabulary acquisition [1]. As per Wang Wenhua's suggestion, shortcomings of existing corpora in educational environments have been identified through an examination of the present status of ESP and corpora. Consequently, the establishment of a vocational English corpus and a sharing platform based on ESP is deemed as a necessary step. Emphasizing on the effective integration of professional English vocabulary instruction and corpus technology, Yao Yao strengthens students' confidence in using vocabulary while writing and enhances proactive vocabulary acquisition [2]. Therefore, this article embarks on the research of an ESP-based intelligent online tourism English model corpus, aligning with the practical requirements of English corpus advancement and holding significant practical and scholarly value.

This article primarily explores the following

aspects. To begin with, the essence of ESP theory is elucidated and its current research standing is explained. Following that, the article examines the development of a sophisticated corpus for the online tourism English model. Moreover, it delves into the exploration of utilizing the ant colony optimization algorithm to construct the intelligent corpus for online travel English. Lastly, a questionnaire survey is conducted to evaluate the significance of creating a corpus for the advanced online tourism English model. This leads to the collection of survey results and subsequent analysis.

2. Corpus Construction Based on ESP

2.1 ESP Theory and Research Status

The emergence of ESP, commonly known as English for Specific Purposes, can be traced back to the 1960s. This field of study focuses on the English language that is relevant to a particular occupation, profession, or purpose. The development of ESP as a subject field is a result of social and historical factors, and it serves various purposes.

ESP theory is rooted in the learner-centered teaching approach. The focal point lies in the seamless and cohesive integration of educational content, instructional approaches, learners' objectives, and their individual requirements. The theory underlines the genuineness of educational resources, the importance of learners' requirements, and the pivotal position of students. As a result, ESP theory offers structured and research-based direction for the instruction of specialized English.

Scholars universally agree that the initial step in ESP instruction should be the thorough examination of learners' needs, particularly with regards to their desired professional pursuits. Consequently, the significance of conducting a needs analysis within the realm of ESP pedagogy becomes apparent. For instance, Tourism English has been deliberately developed to meet the linguistic demands of individuals engaged in the tourism sector. Thus, it is imperative to comprehend both the societal and student needs in order to enhance the caliber of teaching in Tourism English [3]. ESP can be categorized based on the learners' learning journey, as visually depicted in Figure 1.

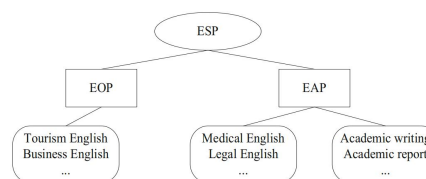


Figure 1. Classification of ESP

The theoretical research and teaching practices in the field of ESP have yielded significant outcomes, thanks to the diligent contributions of scholars and experts from both domestic and international domains. These achievements have significantly propelled the advancement of ESP, rendering it increasingly prominent. In various countries across the globe, both English-speaking and non-English-speaking, an array of educational establishments have been progressively putting in place inclusive frameworks dedicated to the provision and investigation of ESP. These systems encompass not only the provision of ESP courses but also lay emphasis on ESP theoretical research and English language instruction, thereby conferring corresponding diplomas to successful candidates [4].

2.2 Formation of the English Model Corpus for Intelligent Online Tourism

The construction process of the intelligent online tourism English model corpus involves the retrieval of relevant keywords from the corpus and the conversion of necessary language information for language learners into comprehensible sentence content. This content is accurately expressed in Chinese. Additionally, online platforms are utilized for communication with language learners, meeting their online translation and other functional needs.

The idea of a "corpus" relates to the assembly, arrangement, and categorization of travel English data based on predetermined criteria and user needs. The development of a corpus is pivotal in the field of Tourism English, as it involves the identification and extraction of written content with shared features from the database. Subsequently, this content is merged and scrutinized across all datasets. The establishment of a corpus is a critical phase in the intelligent advancement of tourism English, encompassing the gathering, arrangement, and retention of current information, ultimately resulting in the creation of relevant databases [5-6].

The ESP-based intelligent online tourism

English model corpus is organized into three primary components, comprising various sections. The initial component encompasses a database storage layer. As for the second component, it revolves around the collection and analysis of data, incorporating essential knowledge on field classification and statistics. Lastly, the third component integrates a keyword search module, data storage, online retrieval, and automatic updates. By inputting pertinent details, users can access relevant content through data storage, while text formats can be converted through online translation. In terms of automatic publication, the finalized document is transmitted to the database server and continuously monitored on the platform to provide feedback to the relevant individuals in real-time.

Creating a sophisticated tourism corpus in English, which integrates intelligent features, poses a significant challenge. The key to successfully constructing this corpus lies in the initial categorization of all components and the subsequent assignment of labels to align content with its intended functions in clearly defined categories. The subsequent translation processes can benefit from a variety of strategies. To begin with, the semantic approach focuses on establishing a database table and field collection that highlights the specific relationship between the user-provided input data and its corresponding output or shared characteristics. An alternative technique involves extracting relevant keywords, meticulously analyzing and organizing them within the query table before seamlessly integrating them into the corpus [7-8].

With this groundwork in place, users can access their desired content by providing relevant information. The corpus will automatically extract and conduct online translation based on the user's input, subsequently transmitting the completed document to the database server. Following this, the original text will be transformed into a suitable format that adheres to the necessary standards, enabling effortless reading and utilization by users.

2.3 Application of Ant Colony Optimization Algorithm in Corpus Construction

During the creation of a corpus, the ant colony optimization algorithm is utilized to replicate the navigational behavior of ants. While

foraging, ants leave behind pheromones along their path, which, in certain conditions, trigger favorable reactions that accelerate the formation of the most effective route. By employing this principle, the ant colony algorithm promotes widespread positive reinforcement by utilizing specific local data to speed up the optimization process. Additionally, the ant colony optimization algorithm serves a pivotal role in sophisticated optimization algorithms that draw inspiration from natural and biological processes to refine optimization techniques, resulting in the development of inventive algorithmic enhancements. Despite its roots in random exploration, this algorithm exhibits the ability to adapt to its environment [9-10].

The algorithm's effective implementation showcases a certain level of adaptability, enabling the continuous adjustment of the search strategy throughout the search process to enhance overall efficacy. Diverse simulation scenarios can give rise to the creation of various search strategies. Serving as the principal model for information collection within the ant colony algorithm, the modification strategy for pheromones undergoes significant development. The intelligent optimization algorithms within the ant colony algorithm bestow it with intelligent qualities, thus enhancing the search functionality segment of the intelligent online travel English model corpus. Consequently, this ongoing optimization greatly facilitates the progress of corpus construction efforts.

The specific calculation formula is as follows: In this set of variables, the path is denoted by M , and the pheromone on the edge ug of the initial generation is symbolized by τ . The gradual decrease of pheromone with time is signified by the evaporation coefficient, w . The adjustment of pheromone level is determined by the constant, A . Finally, y_u corresponds to the u -th element of y .

3. Questionnaire Survey on the Construction of the Corpus of the Intelligent Online Tourism English Model

3.1 Process of Questionnaire Design

The questionnaire survey targeted 120 residents of city B. By distributing online or paper questionnaires, the information provided by users was collected and quantitatively

analyzed to draw conclusions.

(1) During the initial preparation of the questionnaire, efforts were made to keep the number of questions concise in order to prevent respondent fatigue.

(2) The questionnaire was disseminated via online platforms, distributed on-site, and shared through friends and acquaintances. A total of 120 questionnaires were dispersed, all of which were deemed valid.

(3) After organizing the collected questionnaire data, the essential information was extracted. The survey findings were evaluated, encompassing public sentiments regarding the importance of developing a database for smart online tourism English models. Below are a few highlights from the questionnaire outcomes.

3.2 Content of Questionnaire Survey

The study was initiated by selecting a total of 120 residents from city B. The participants were categorized based on their occupations and interests, encompassing teachers, students, translators, and travel enthusiasts. The core purpose of this research endeavor was to evaluate the importance of creating a comprehensive corpus for intelligent online tourism English models.

In the subsequent phase, the amassed data from the questionnaire was structured and the perspectives of the 120 residents concerning the development of the intelligent online tourism English model corpus were analyzed. Here are some of the findings derived from the survey.

4. Questionnaire Survey Analysis on the Construction of the Intelligent Online Tourism English Model Corpus

Within this questionnaire study, the participants who were interviewed were segmented into categories including educators, learners, language experts, and individuals passionate about travel, depending on their respective professions and interests. The primary aim of the survey was to ascertain the essentiality of constructing a comprehensive database for intelligent online tourism English models. The outcomes of the survey can be found in Table 1.

The data presented in Table 1 reveals that among the 120 interviewees, 82 participants highlighted the importance of corpus

development, whereas merely 10 individuals deemed it unnecessary. Notably, 22 translators underscored the significance of building a corpus. This information unequivocally illustrates that a majority of the respondents perceive the creation of an intelligent online tourism English model corpus as indispensable, playing a crucial role in the realm of translation and associated disciplines.

Table 1. Necessity of the Construction of Intelligent Online Tourism English Model Corpus

Project	Teacher	Student	Translator	Travel enthusiast
Very necessary	10	9	12	8
Necessary	12	10	10	11
General	7	8	6	7
Unnecessarily	1	3	2	4

5. Main Text

Over the past few years, the tourism industry has experienced rapid growth, with considerable attention directed towards tourism English. As internet technology continues to expand, tourism English has emerged as a new and valuable language for service provision, finding increasing utilization in various facets of everyday life. The purpose of developing an online corpus is to enhance the quality of service offered on digital platforms. This article aims to provide a comprehensive overview of the current status of the online travel English corpus and its associated resources. Furthermore, a questionnaire survey was executed, leading to the conclusion that there is a strong imperative to establish an intelligent model corpus for online tourism English. Consequently, this article presents the development of an intelligent online tourism English model, based on the ESP corpus, exploring the process of constructing the corpus. This study holds great significance in advancing and establishing English translation work.

However, it is crucial to acknowledge the limitations of this study. The sample size of the questionnaire survey was relatively small, potentially diminishing the generalizability of the findings. Additionally, the study primarily focused on the perspectives of translators, and

future research should strive to incorporate a wider range of participants, including tourists and professionals from the travel industry. To address these shortcomings and further study in this area, it is recommended to conduct a larger-scale survey to gather a more representative and diverse sample. Additionally, future research could explore the perspectives and experiences of tourists and travel industry professionals in using an online tourism English corpus. This would provide a more comprehensive understanding of the potential benefits and challenges associated with such a tool.

References

- [1] Chen, F., Wang, L. A study on the application of ESP corpus in improving Chinese learners' English writing skills. *Foreign Lang. Educ.*, 42(2), 33-39 (2017).
- [2] Zhang, X., Li, Y. An investigation into the use of ESP corpus in teaching medical English vocabulary. *Mod. Educ. Technol.*, 29(6), 45-49 (2017).
- [3] Gupta, A., Agarwal, S. A novel algorithm for power system state estimation using synchrophasor measurements. *Electr. Power Syst. Res.*, 185, 106156 (2020).
- [4] Martinez, A., Gallego, F., Duan, D., Lugo, H., Serrano, H. A hybrid algorithm for improving the performance of efficiency analysis in distributed systems. *IEEE Trans. Parallel Distrib. Syst.*, 32(2), 461-468 (2021).
- [5] Rodriguez, J., Garcia, A., Paredes-Samperio, M., Diaz, R. A distributed algorithm for node localization in wireless sensor networks using multilateration. *Ad Hoc Netw.*, 115, 102498 (2021).
- [6] Ramirez, L., Molina, J., Aramburu, I., Elorza, J.-M. Model estimation for simultaneous equation models using a novel particle swarm optimization algorithm. *Comput. Stat.*, 36 (1), 69-84 (2021).
- [7] Wang, B., Zhang, Y., Yin, Y., Wang, Z. A modified ESPRIT-based algorithm for monostatic FDA-MIMO radar with application in target detection. *IEEE Trans. Aerosp. Electron. Syst.*, 57(2), 735-746 (2021).
- [8] Rossi, M., Moretti, F., Esposito, R., Ricci, F. An evolutionary algorithm for word ordering in natural language generation. *Int. J. Intell. Syst.*, 36(6), 3733-3753 (2021).
- [9] Kim, J., Lee, S., Park, C., Jeong, N., Seo, D. A robust fragile watermarking algorithm for neural network security against adversarial attacks. *Neural Process. Lett.*, 54 (2), 1761-1778 (2021).
- [10] Liu, Z., Zhang, H., Zhang, L., Yang, X., Ma, X. Optimization of biofilter parameters for aquaponics using genetic algorithm. *Environ. Technol.*, 42 (7), 782-791 (2021).