Construction of Business English Needs Analysis Model Based on Computer Corpus

Yao Lu[1(✉)], Li Zhang[1,2], and Lingya Ma[2]

[1] Shandong Management University, Jinan 250357, Shandong, China
465666751@qq.com
[2] Hubei Three Gorges Polytechnic, Yichang 443000, Hubei, China

Abstract. In order to meet the needs of business English, this article constructs a computer corpus based business English requirement analysis model. By collecting and organizing relevant corpus in the field of business English, establish a large-scale and diverse corpus. Then, using natural language processing technology and machine learning algorithms, the corpus is analyzed and processed to extract common vocabulary, phrases, and grammatical structures in business English. Next, by clustering and mining association rules on the texts in the corpus, the actual needs and characteristics of business English are discovered. In the process of model construction, we will also consider different branches and specialties in the business field, such as marketing, international trade, business communication, etc., to adapt to the needs of different fields. Finally, based on the results of the needs analysis, we will provide targeted suggestions and guidance for business English teaching and training. The construction of this model helps educators and learners better understand the characteristics and needs of business English, in order to carry out business English teaching and training more effectively.

Keywords: Corpus · business English · big data · computer

1 Introduction

With the continuous development of global business, the demand for business English is growing day by day. However, the characteristics and needs of business English are not entirely the same as general English, so customized teaching and training are needed to meet the special needs of the business field [1]. In order to better meet the needs of business English learners, this article aims to construct a computer corpus based business English needs analysis model.

This study will utilize a large-scale computer corpus that includes relevant texts in the field of business English [2]. We will analyze and process the corpus through natural language processing technology and machine learning algorithms to extract common vocabulary, phrases, and grammatical structures in business English [3]. By using data mining techniques such as clustering and association rule mining, we will explore the practical needs and characteristics of Business English.

Y. Zhang and N. Shah (Eds.): BigIoT-EDU 2023, LNICST 580, pp. 496–503, 2024.
https://doi.org/10.1007/978-3-031-63130-6_55

At the same time, we will also consider different branches and majors in the business field, such as marketing, international trade, business communication, etc., to meet the needs of learners in different fields [4]. Through the results of demand analysis, we will provide targeted suggestions and guidance for business English teaching and training, helping learners better cope with language communication challenges in business scenarios [5].

The results of this study are of great significance for the development of business English education, as they can improve learners' business English proficiency and promote their language application abilities in business environments [6]. At the same time, the construction of this model also helps educators better understand the characteristics and needs of business English, thereby providing more effective teaching strategies and resources [7].

2 Construction Principles of English Corpus

Building an English corpus requires following certain principles to ensure the quality and applicability of the corpus. Firstly, the corpus should include various types of text, such as news reports, scientific papers, novels, blog articles, etc. This can cover different fields and language styles, making the corpus more representative and practical.

Secondly, corpora should cover different language levels and difficulty levels to meet the needs of different learners. This means that it should include basic, intermediate, and advanced language materials, allowing learners to gradually improve their English proficiency [8].

In addition, the content of the corpus should have timeliness, including the latest news reports, research results, and social media content. This helps learners to understand current language usage and popular topics, and improve their English practical skills.

In addition, the diversity of corpora is also crucial. The corpus should include various themes and genres, covering economic, political, technological, cultural, and other aspects. This can enable learners to find materials of interest in different fields and enrich their vocabulary and knowledge reserves [9].

Finally, the corpus should be large-scale to provide sufficient samples to support language analysis and research. This means collecting and organizing a large amount of textual data, including books, newspapers, web pages, etc. At the same time, attention should also be paid to protecting the copyright and privacy of the corpus, and ensuring that the corpus is obtained and used in a legal and ethical manner.

In short, building an English corpus requires consideration of principles such as diversity, representativeness, timeliness, and scale, in order to meet the needs of learners and promote the research and development of the English language. As an alignment positional relationship that can be represented by byte offsets, XML documents can identify corresponding positions based on each keyword field and specify main information such as chapters, paragraphs, sentences, and words. The XML annotation set can be divided into chapters, paragraphs, sentences, and words, as shown in Fig. 1.

Constructivism is an educational and learning theory that believes that knowledge is constructed through the interaction and communication between individuals and the environment. According to constructivism, learning is not only about receiving and

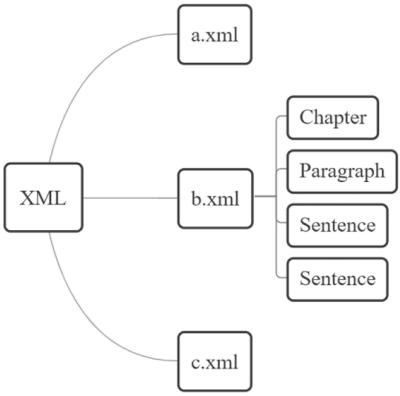

Fig. 1. Corpus storage and annotation function

storing information, but more importantly, individuals actively construct and understand knowledge [10]. Constructivism emphasizes learners' initiative and participation. It believes that during the learning process, individuals construct their own understanding and meaning through interaction and social interaction with the real world. Learners construct new knowledge and concepts through the process of practice, exploration, and problem-solving. Constructivism also emphasizes the social nature of learning. It believes that learning is carried out in social interaction, where individuals can jointly construct and share knowledge through collaboration and communication with others. By collaborating, discussing, and reflecting with others, learners can expand their cognition and understanding by observing their behavior and listening to their perspectives.

In addition, constructivism also emphasizes the influence of individual backgrounds and experiences on knowledge construction. Each learner has their own unique background and experience, which can affect their understanding and construction of knowledge. Therefore, educators need to pay attention to the individual differences of learners and design and support teaching based on their backgrounds and needs.

As shown in Fig. 2, the lack of a large number of vivid practical situations in traditional business English requirements has had a negative impact on the construction of student needs.

Foreign language circles have put forward various demand analysis models or frameworks, but most of them are aimed at the demand for business English or business English for special purposes. There is a lack of demand analysis and Research on the demand for business English for English majors [11]. As a new interdisciplinary subject, business English has a strong development momentum. In better the demand for business English, It is for us to explore and construct a business English needs analysis model based on the foreign language needs analysis model.

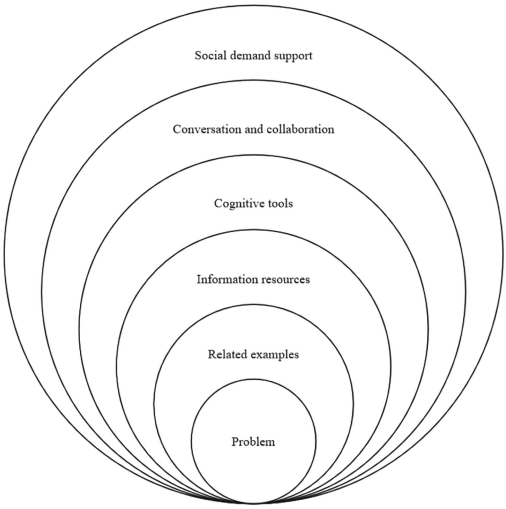

Fig. 2. Business English demand environment model

3 Keywords and Criticality

It is possible to extract some corpus keywords and calculate the key of keywords in the corpus. However, it is not scientific and accurate to calculate keywords and their key only based on the corpus. For example, if we only refer to the frequency of words in this corpus, the frequency of some high-frequency words, such as articles and some common conjunctions and prepositions, is often quite high. However, it is inaccurate to determine these words as keywords. Therefore, we must compare the of this word corpus [12].

The processing and identification process of different information is shown in Fig. 3.

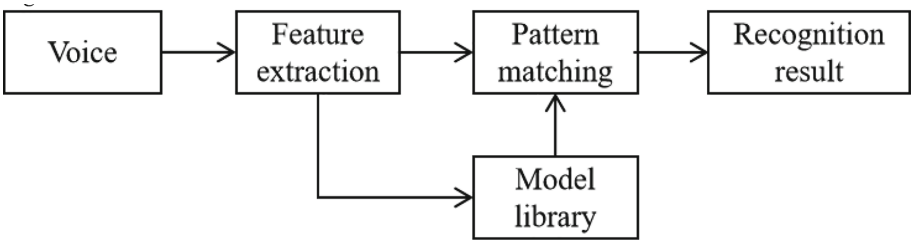

Fig. 3. Information processing and identification process

Generally speaking, the have a large scale, at least significantly larger than the ship model of the studied corpus, so as to play an role [13]. If the frequency of a corpus is not high, but it is relatively high in the studied corpus, it can be preliminarily determined that the word is a keyword in the studied corpus. The key of keywords can be by x2 value. The menu word in the studied, the length of the studied corpus is m in the number of glyphs, the frequency of the length of the reference corpus is n in the number of glyphs,

then (1):

$$x^2 = (|fn - cm| - (f + c + m + n)/2)(f + c + m + n)/ \\ (f + n)(f + m)(f + c)(m + n)) \tag{1}$$

It is generally believed that if the X2 value of a word, it has key significance in the studied corpus, that is, it has obvious key. The determination and calculation of the keywords and keys of the studied corpus have great potential in the Business English needs [14]. Using this technology, it is easy to automatically extract keywords, which can help the subject classification of business English needs.

4 Construction of Business English Demand Model Based on Corpus

4.1 Construction of Industry Demand Analysis Model in the Construction of English Major

Business English is an important branch of English for special purposes (ESP), and needs analysis is its remarkable feature. ESP teaching has many problems in curriculum positioning, curriculum standards, teaching modes and methods. This is incompatible with the current social needs of employment objectives and the basic requirements of the teaching objectives of Liang. In order to clearly define these goals, it is necessary to analyze the needs of business English workers and future goal scenarios. The future target scenario of business English major is the front-line foreign-related posts in foreign trade enterprises and institutions [15]. Therefore, the demand analysis of business English industry is to analyze the English language knowledge and skills, business knowledge and skills and professional quality of front-line foreign-related posts in foreign trade enterprises and institutions, so as to closely combine with the work process and work needs of business English workers, and provide a scientific basis for the construction of business English major [16]. As shown in Fig. 4.

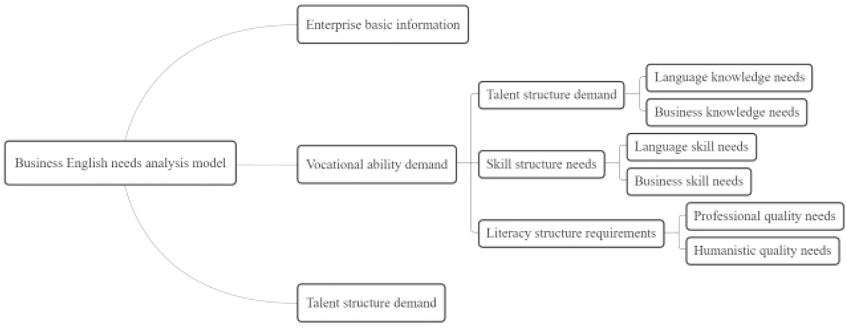

Fig. 4. Business English demand analysis model

4.2 Intelligent Screening of Japanese Network Corpus Information Based on Corpus

The refers to the process of using artificial intelligence technology and corpus resources to intelligently screen and extract useful information from Japanese network corpus. This intelligent filtering can greatly improve the efficiency and accuracy of Japanese learners and researchers in obtaining effective corpus in the online environment.

Assuming that the property library entry T, and T appears the same number of times in each document. As shown in formula (2).

$$w_{itk}^{(s)} = \frac{tf_{itk}}{l_{it}} \tag{2}$$

The of Japanese online on corpus requires the construction of a rich and diverse Japanese online corpus. This can include various types of online text, such as news reports, blog articles, social media content, etc. Such a corpus can cover different themes and language styles, meeting the various needs of learners and researchers. If the shown in formula (3).

$$w_{jk}^{(q)} = \begin{cases} 1, T_k \in S_{it} \wedge T_k \in Q \\ 0, T_k \in S_{it} \wedge T_k \notin Q \end{cases} \tag{3}$$

The above formula and in combination with formula (4)

$$S(q_j, s_{it}) = \frac{\sum_{k=1}^{n} w_{jk}^{(q)} w_{itk}^{(s)}}{\sum_{k=1}^{n} (w_{jk}^{(q)})^2 \sum_{k=1}^{n} (w_{itk}^{(s)})^2 - \sum_{k=1}^{n} w_{jk}^{(q)} w_{itk}^{(s)}} \tag{4}$$

The screening of Japanese online can also apply language models and machine translation technologies to provide auxiliary functions for language learning and understanding. By analyzing and understanding language patterns and syntactic structures in the corpus, intelligent screening systems can provide learners with relevant language knowledge such as example sentences, phrase combinations, and grammar explanations, helping them better understand and apply Japanese. Formula (5) is:

$$S(q_j, d_i) = \frac{1}{N} \sum_{t=1}^{N} S(q_j, s_{it}) \times \eta_t \tag{5}$$

By utilizing artificial intelligence technology and rich corpus resources, useful Japanese online corpora can be efficiently extracted and screened, providing personalized learning and understanding assistance. Such an intelligent screening system can help learners and researchers better master and use the Japanese language. Based, obtain the formula for the model of business English demand corpus information (6)

$$B_{ij} = s\left(q_j, d_i\right) \times \sum w_{jk}^{(q)} \times s(i, j) \tag{6}$$

Overall, the utilizes artificial intelligence technology and natural language processing technology to help users efficiently screen and extract valuable information from the corpus, improving the efficiency and accuracy of language learning and research.

5 Conclusions

This paper introduces the theoretical basis of corpus based auxiliary English demand, and preliminarily discusses the classification of demand corpus and the of different content in English demand caused by the idea of classification. When analyzing English needs, it can be found that the relevant teaching concepts embodied in communicative approach can be reflected to a certain extent [17]. It believes that language is the main carrier of communication and the means for people to actively communicate in today's society.

This paper mainly combines the above-mentioned relevant models of demand analysis to conduct a positive demand analysis on business English, and establishes the analysis model. Generally speaking, it mainly includes target demand analysis and current demand analysis. The target scenario analysis mainly includes four dimensions: Business English information, business English communication information The basic situation and the adaptability of business English to social needs [18]; Generally speaking, when analyzing the current English needs, we need to consider four contents, namely, business English work needs, business English language information and business environment needs.

References

1. Chen, B.: Construction of industry demand analysis model in the construction of business English major in Higher Vocational Colleges **2017**(1), 121–124 (2020)
2. Zhang, Y., Jing, Y.: Construction of business English corpus based on foreign economic needs – taking tea trade English corpus as an example. Fujian Tea **40**(193), 243–244 (2018)
3. Xiao, L., Zhou, M., Yan, M.: Demand analysis and capability model construction of business English talents in the Yangtze River Delta **2020**(8), 165–166 (2021)
4. Hong, W.: A corpus based multi-dimensional comparative analysis of business English letters. J. Huaihai Inst. Technol. Humanit. Soc. Sci. Edn. **14**(3), 5 (2021)
5. Zheng, J., Wu, T.: Visual analysis of domestic ESP teaching research trends – based on the survey of 11 foreign language core journals (2007–2017). Hebei Agric. J. (Soc. Sci. Edn.) **1**(201), 54 (2019)
6. Wu, J.: Analysis of differences between business English and general English corpora based on multi digit analysis. Modern Vocat. Educ. **000**(31), 176 (2018)
7. Jiang, H.: Analysis of corpus based project-based teaching reform of business English interpretation. Hubei Correspondence J. **031**(012), 162–164 (2018)
8. Zhang, S.: Exploring key elements in the construction of business English linguistics. J. Jiamusi Inst. Educ. (007), 418 (2021)
9. Wang, N.: Construction of logistics demand combination improvement and prediction model based on IOWHA method. Bus. Econ. Res. **14**(1), 67–68 (2017)
10. Wu, X.: An analysis of the project-based teaching reform of business English interpretation based on corpus. J. Liuzhou Teachers College **32**(6), 64–66, 63 (2017)
11. Cao, B.: Demand analysis of business English Majors Based on regional economic development. Friends of middle school English: Research on foreign language learning and teaching methods **2**(1), 272 (2017)
12. Jiang, J., Xu, J.: Multidimensional analysis of register features of business English based on corpus **2015**(2), 225–236 (2021)

13. Zhang, S.: Exploring several key elements in the process of Constructing Business English linguistics. J. Jiamusi Vocational College (7), 418 (2017)
14. Xie, L.: Research on business English teaching from the perspective of corpus – Taking the cultivation of export-oriented tourism talents as an example. Education: Higher Educ. Res. (12), 43–44 (2021)
15. Wang, C.: Research on the construction of competency training system model for business English majors from the perspective of employability. J. Chengdu Normal Univ. **34**(5), 18–23 (2018)
16. Xu, X.: Building a training model for business English talents based on demand analysis **6**, 31–32 (2018)
17. Si, L., Shi, Y.: Research on the construction method of comparable corpus based on cross language information retrieval **2016**(6), 64–70 (2018)
18. Xiao, L., Zhou, M., Yan, M.: Demand analysis and competency model construction of business English talents in the Yangtze River Delta. Modern Commun. 526(8), 169–170 (2018)