



Evolutionary Trends and Opportunity Identification of University English Learning Patterns Based on Louvain Clustering Algorithm

Fang Xia¹, Xin Liu¹, and Li Yin²(✉)

¹ School of Health Management, Changchun University of Chinese Medicine,
Changchun 130117, China

² School of English Teaching Department, Changchun University of Chinese Medicine,
Changchun 130117, China
53649072@qq.com

Abstract. This study aimed to analyze the current state of research on college English learning models and to identify opportunities in this area by analyzing evolutionary trends. At present, the demand for university English learning mode research is expanding, and the university English learning mode research literature is sorted out with the introduction of clustering algorithm and the effect of opportunity identification. In order to gain a deeper understanding of the research progress of English learning models in universities, the article uses the Louvain clustering method to map the number of articles, research institutions, keywords and hotspots in the research field by using 1866 documents related to English learning models collected on the Internet, analyzing their distribution patterns and evolutionary trends and identifying opportunities for future research directions. The results show that the academic attention of college English learning mode climbs year by year, the institutions in the field do not cooperate closely, the blended learning mode is the research hotspot of college English learning mode, and the blended English teaching mode based on big data is the frontier area of college English learning mode research.

Keywords: College English learning patterns · Louvain clustering algorithm · Research hotspots · Evolutionary trajectories · Opportunity identification

1 Introduction

In the past forty years, university English education in China has made great progress and development, making great contributions to the country's economic development and cultural construction. However, with the new situation and new requirements emerging from higher education in the new era, the model of university English learning is also facing considerable controversy and crisis [1]. Accordingly, on September 6, 2022, the website of the Ministry of Education published the Teaching Requirements for College English Courses, which stated that the design of college English courses should make extensive use of advanced information technology and develop and build a variety of computer and web-based courses to provide students with a good language environment.

© The Author(s) 2024

F. Huang et al. (Eds.): ICAIE 2023, AHCS 15, pp. 611–619, 2024.

https://doi.org/10.2991/978-94-6463-242-2_75

The intervention of computers and other big data tools has made the research and construction of English learning models a hot topic in major universities. Scholars in China have made fruitful theoretical research results in the field of college English, which have been widely used in college English classroom teaching practice. However, the existing research results also have problems such as repetitive research and outdated research questions [2]. This paper analyzes 1866 documents collected on the Internet based on the Louvain clustering algorithm, with the aim of discovering the development history, evolutionary trajectory, research frontiers, and opportunity identification of college English research in China in the past decade, helping researchers in the field to keep abreast of relevant research developments, avoiding duplication of research, and providing direction for further research.

2 Data Sources and Research Algorithms

2.1 Data Source

In order to comprehensively analyze the current situation and evolutionary trends of university English learning models, the article uses the CNKI database of China Knowledge Network as the data source, selects the “advanced search” type, and searches with “university English learning models” in the title of the article, and the matching method is “The search was conducted on November 10, 2022.” A total of 3150 documents were retrieved, and the documents were merged, examined, deleted, extracted, and de-duplicated, excluding conference abstracts, journal calls for papers, news, information, reports, etc. A total of 1866 relevant documents were selected as the data source of the articles.

2.2 Research Algorithms

Louvain clustering algorithm, Louvain algorithm is a modularity-based community discovery algorithm. The basic idea is that nodes in the network try to traverse the community labels of all their neighbors and select the community label that maximizes the modularity increment. After maximizing the modularity, each community is viewed as a new node and repeated until the modularity is no longer increasing [3].

Modularity algorithm:

$$Q = \sum_c \left[\frac{\sum in}{2m} - \left(\frac{\sum tot}{2m} \right)^2 \right] \quad (1)$$

where $\sum in$ denotes the weights inside the community c , and $\sum tot$ denotes the weights of the edges connected to the points inside the community c , including the edges inside the community as well as the edges outside the community. The modularity Q gain obtained in moving to a community C can be easily calculated as:

$$\begin{aligned}\Delta Q &= \left[\frac{\sum in + K_{i,in}}{2m} - \left(\frac{\sum tot + K_i}{2m} \right)^2 \right] - \left[\frac{\sum in}{2m} - \left(\frac{\sum tot}{2m} \right)^2 - \left(\frac{K_i}{2m} \right)^2 \right] \\ &= \frac{1}{2m} \left(K_{i,in} - \frac{\sum tot K_i}{m} \right)\end{aligned}\quad (2)$$

where $\sum in$ is the sum of the weights on the links within the community C (when the weight is 1 it is equal to the degree), if it is the initial case, i.e., when a node as a community, it is the connection from this node itself to itself, which still requires the starting point plus weight and the end point plus weight (even if the starting point and the end point are the same node at this time, for the acyclic graph weight is 0, the $\sum tot$ is the sum of the weights on the links associated to the nodes in C is the sum of the weights of the links associated to node i, $K_{i,in}$ is the sum of the links from node i to the nodes in C, and m is the sum of the weights of all the links in the network.

The evolutionary trends of keyword clustering, authors, institutions, and hotspots of university English learning mode research were processed and mapped based on the Louvain clustering algorithm for the literature sample to explore the evolutionary trends of research in the field of university English learning mode.

3 The Evolutionary Trajectory of University English Learning Model Research

3.1 Evolution of the Number of Articles Issued

The number of annual publications is an important criterion to measure the hotness of research on English learning models in universities, and the analysis of 1866 papers published in the field of English learning models in universities from 2012 to 2022 shows that the number of publications from 2012 to 2015 is 259, 231, 272 and 279. The number of articles peaked in 2015, indicating that the research on English learning mode in universities is in a rapid development stage, which is mainly benefited from the concept of "Internet+" first proposed in 2012 and widely used in the field of English learning by researchers in the following years. The research trend is on the rise from 2019 to 2021, with the national health environment influencing the learning mode of English in universities, and most universities starting to adopt online teaching and learning mode, with the number of articles published reaching 259 in 2019. This trend shows that with the change of research environment, the research on English learning mode is also developing in depth, reflecting from the side that there is still more space in the field of college English learning mode. As shown in Fig. 1.

3.2 Distribution of Major Research Institutions

The analysis of the cooperation of the issuing institutions shows that there are 25 research institutions with the total number of articles greater than 5 on college English learning mode from 2012 to 2022, among which the institution with the most articles is the

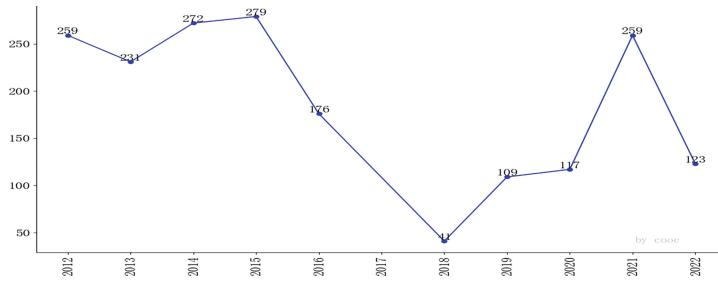


Fig. 1. Annual number of articles published in the field of English learning model research

Department of Foreign Languages of Harbin Engineering University, which has earlier research on college English learning mode and rich research experience, and the main direction is the improvement of English listening ability. The main direction is the improvement of English listening ability. The difference in the total number of articles issued by other institutions is not very large, and universities account for most of the research institutions, but the similarity of the research samples is too high, which leads to the lack of close cooperation among universities. See Fig. 2.

Liaodong University is second only to Harbin Institute of Technology in terms of the number of articles published, and its research direction is the “3 + 1 + X” English learning mode; the research progress of different regions also varies, with Nantong University’s research topic being online independent learning, and the research on online learning has been conducted earlier than most institutions. This indicates that the field of English learning models should continue to find breakthrough points in hot topics, link samples from different universities to conduct research, enhance the intensity of cooperation among institutions, and strengthen cross-regional and cross-disciplinary cooperation among research institutions.

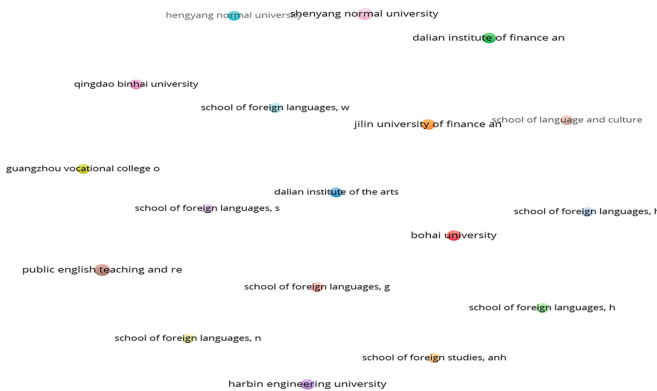


Fig. 2. Distribution of major institutions in the field of university English learning model research

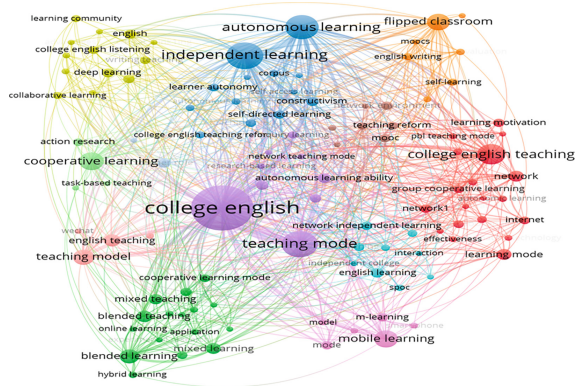


Fig. 3. Co-occurrence mapping of major keywords in the field of university English learning mode research

4 The Evolutionary Trajectory of the Research Hotspot of the University English Learning Model Collar

Research hotspots can reflect the development pulse and direction of research in an academic field within a certain period of time [4]. The article analyzes the general research hotspots in the field by mapping the co-occurrence of high-frequency keywords in the field of university English learning mode research through Louvain's algorithm, followed by removing the less meaningful high-frequency keywords and mapping the word cloud to analyze the specific research hotspots in university English learning mode.

4.1 Keyword Co-occurrence Analysis of the Study of University English Learning Mode

Firstly, keywords were extracted from 1866 documents in the data sources, and a total of 2291 keywords were obtained and calculated by Louvain's algorithm. The higher the frequency of the keywords, the more obvious the centrality of the point degree in its algorithm, which indicates that the keywords are more important in this research field. The comprehensive analysis of Fig. 4 shows that the nodes and centrality of these keywords are larger and constitute the basic research direction of the university English learning mode, which is closely associated with the research direction and constitutes the basic elements of the university English learning mode research. The other high-frequency keywords besides the research themes are independent learning, cooperative learning, teaching mode, and blended English learning mode. The high frequency and centrality of these keywords can be judged as the basic research areas of university English learning mode (Fig. 3).

4.2 Analysis of Specific Research Hotspots in the Field of English Learning Models in Universities

Based on the analysis of the general research hotspots of college English learning mode, the analysis of specific research hotspots removed some high-frequency keywords that

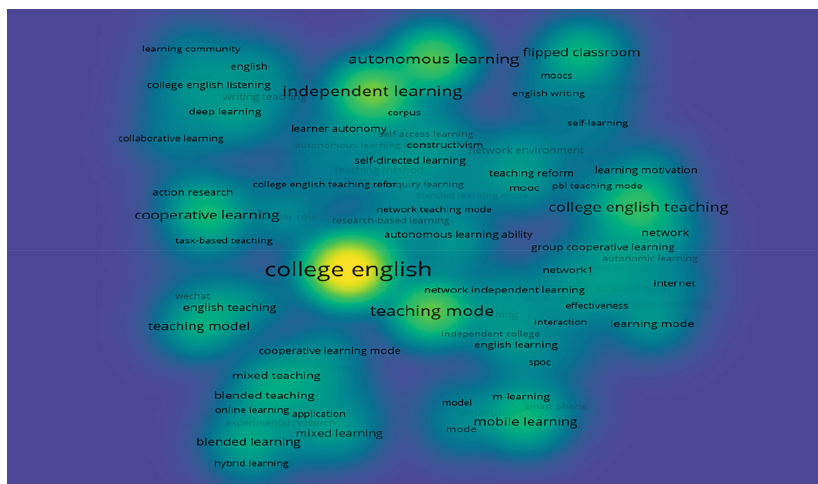


Fig. 4. Keyword density in the field of English learning mode research at university

do not reflect specific research hotspots, such as college English, learning mode, teaching mode, etc., and drew a density map of keywords in the research area of college English learning mode from 2012 to 2022, so as to simply and clearly see the learning models. This is shown in Fig. 4.

It can be concluded from Fig. 4 that the specific research hotspots in the field of college English learning mode research from 2012 to 2022 can be summarized into three research areas: students' own aspects, teaching tools, and learning platforms. The research on students' own aspect is mainly manifested in the keywords of independent learning, cooperative learning, online independent learning, mobile learning, writing ability, listening ability, etc.; the research on teaching methods is mainly manifested in the keywords of flipped classroom, constructivism, teaching reform, blended teaching, teacher's role, etc.; the research on learning platform is mainly manifested in the keywords of online environment, online teaching, online platform, etc. Since the learning of college English has an important role in professional English learning, English literature reading and work, many scholars began to focus on the learning mode, teaching mode and learning platform of college English.

4.3 Keyword Evolutionary Trajectory and Opportunity Identification of University English Learning Model Research

Two indicators, the number of articles and the frequency of keywords in university English learning mode from 2012 to 2022, were incorporated into the Louvain algorithm for calculation and plotted into a weighted time-zone map of keyword evolution year by year to visualize the keyword evolution trend in the field of university English learning mode research.

Figure 6 shows that the evolutionary development vein of high-frequency keywords of university English learning mode is roughly divided into three stages: the first stage

is 2012–2016, which is a period of rapid development of university English learning mode research and has made great progress; the second stage is 2017–2019, this stage of university English The field of learning models entered a phase of calm development, but the innovation of English learning model research was strengthened, and the hotspots of keywords such as mobile learning and flipped classroom rose; the third phase is 2019–present, and the keywords emphasize blended learning models, which are a combination of traditional teaching and online teaching, and require the integrated use of different learning theories, technologies and means and applications to provide students with optimal delivery methods to obtain the most effective language learning practices [5].

As shown in Fig. 5 English learning models are being explored in research on blended model learning models, and in conjunction with changes in the educational environment it can be determined that English learning models will evolve in the future toward blended English learning models that combine with educational big data.

Big data in education is a brand news field. Foreign research on big data in education has been conducted for decades [6] and the field of education technology has created a boom in research related to the development of big data-based technologies for educational reform and innovation. Big Data in education will data mine and analyze the large-scale data generated by students during their English learning and education to help make online decisions on English learning patterns, achieve predictive analysis of students' learning patterns, academic analysis, and industry analysis. The blended English learning model based on big data in education builds on the “Internet+” to build a learning process that combines the digital, online classroom with the traditional classroom.

The big data model and the data shown will provide a reference for students and teachers and will enable timely and accurate assessment of students' English learning status, identify potential problems, and then predict and develop personalized English learning programs. The teacher's role is no longer just to impart knowledge, but to guide,

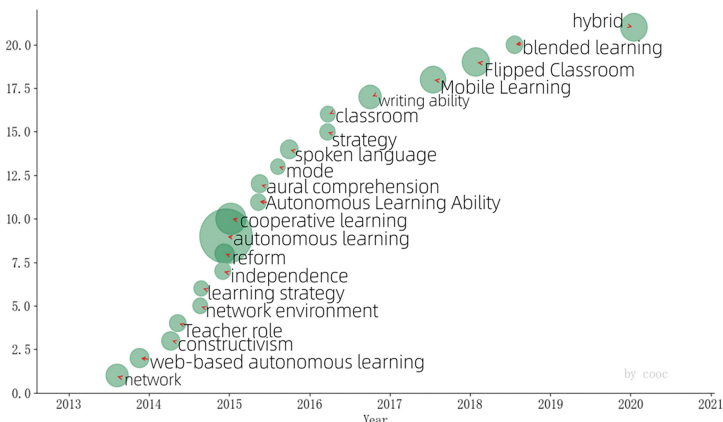


Fig. 5. Weighted time zone diagram of the evolution of research themes in English language learning models in college

inspire and monitor students' learning, focusing on students' achievement of personalized learning goals and their personal emotional experience, so that they can achieve the best learning results [7]. As an international discipline, the choice of teaching and learning mode of English should change with the progress of the times. In the context of today's Internet big data, the teaching mode of English in college has become inseparable from the support of network technology and multimedia technology. The blended college English learning mode based on educational big data focuses on the interactivity among students, emphasizes the centrality of students, advocates a personalized teaching mode, continuously integrates and optimizes online and offline teaching and learning means, guides students to play the main role, maximizes students' enthusiasm and creativity in learning English, and finally realizes the steady state development of the blended English learning mode.

5 Conclusion

This study visualizes 1866 papers published in the CNKI database on university English learning models from 2012–2022 using the Louvain clustering algorithm, focusing on the annual volume of publications, institutional collaboration, keywords, and evolutionary trajectories in the field, predicting opportunities for research and identifying them, and concluding that, first, the university The volume of publications in the field of English learning model research is influenced by the external environment and fluctuates. The institutions with high publication volume and academic influence are mainly gathered in the northeast region; second, inter-institutional cooperation is not close, and cross-regional, cross-discipline and cross-institutional cooperation is less; third, topics such as independent learning, cooperative learning, teaching models and flipped classroom are hot topics of university English learning models; fourth, it is found that the research trajectory of the field has entered into a hybrid English learning model; fifth, it is identified that future English learning models usher in opportunities for the application and research of blended English learning models based on big data.

In general, the field of college English learning mode research has a large research space, and various institutions, authors, and universities still need to invest a lot of time and energy in in-depth research. In the field of foreign language education, there is less relevant research. The research on college English learning mode should start from the teaching goal of diversified needs, explore the teaching method of streaming and sub-module through education big data, stimulate students' independent learning ability, develop personalized and diversified education and learning methods, and focus on the cultivation of students' comprehensive foreign language ability, so that the teaching of foreign language courses can not only meet the effective communication but also improve comprehensive literacy to meet the actual needs of talent cultivation [8].

References

1. Mao W. The relationship between educational policies and the development of English research in China's universities--a knowledge mapping analysis based on CiteSpace [J]. *Journal of Foreign Languages*,2022(03): 92–97. <https://doi.org/10.16263/j.cnki.23-1071/h.2022.03.015>.

2. Cheng, Jing-Yan. A study on the value orientation of foreign language education policy in China in the context of the Guide[J]. *Journal of Foreign Languages*,2021(02): 121–125. <https://doi.org/10.16263/j.cnki.23-1071/h.2021.02.020>.
3. Xue Qing. Research and application of improved Louvain association discovery algorithm[D]. Shanxi University, 2020. <https://doi.org/10.27284/d.cnki.gsxiu.2020.001920>.
4. Wang FQ, Han YP, Wang P, et al. Keyword co-word clustering analysis of water resources research [J]. *Advances in Water Resources and Hydropower Science and Technology*, 2014, 34(4): 29–34.
5. SHARMA P. Blended Learning [J]. *ELT J*, 2010, 64 (4): 456-458
6. Li Yinling, Guo Qiang, Feng Ruiling. A literature review on the study of college English learners' learning behavior in the era of big data[J]. *Educational Modernization*,2019,6(88):240–242+263. <https://doi.org/10.16541/j.cnki.2095-8420.2019.88.094>.
7. Pan Hong. The construction of hybrid teaching mode of “online + offline” English translation theory and practice course under the perspective of “Internet+” [J]. *Campus English*, 2023(05):39–41.
8. Pan J. Research on the teaching strategy of college English based on students' needs analysis-Economics and management majors as an example[J]. *Modern English*,2022(15):13–16.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

