WHIOCE PUBLISHING PTE. LTD.
PROVIDING
FIRST-CLASS SCIENTIFIC INFORMATION
FOR TOP SCHOLARS

2025 Volume 3, Issue 4
Contemporary Education Frontiers
ISSN (Online): 3029-1860 ISSN (Print): 3029-1879

Topic Analysis of English Oral Competence Research Based on LDA Model

Xiaohan Dou, Yun Li*
Shandong University of Science and Technology, Qingdao 266590, Shandong, China

Corresponding author: Yun Li, liyun@sdust.edu.cn

Abstract: This study utilizes the Latent Dirichlet Allocation (LDA) topic modeling method to systematically analyze academic articles related to "English oral competence" published in core domestic journals between 1999 and 2024. By exploring a large corpus of literature, the research identifies three primary themes: English oral competence pedagogy, English oral competence testing and assessment, and English oral competence learning and language analysis. Through a detailed analysis of the evolution of these themes, the study reveals that the application of intelligent technologies, such as artificial intelligence and automated speech recognition, is gradually becoming a focal point in English oral competence research. At the same time, the modernization of oral testing and pedagogy remains a key area of concern for educators and researchers. Future research trends are likely to prioritize the integration of cutting-edge technologies, such as machine learning and AI, alongside interdisciplinary approaches, providing new directions and momentum for the further development of English oral competence research. This study contributes valuable insights into the ongoing transformation of the field.

Keywords: LDA model; English oral competence; Topic mining

Online publication: May 26, 2025

1. Introduction

As one of the core competencies in foreign language learning, second language (English) speaking proficiency has long been a focus of language teaching, as well as one of the key challenges faced by learners[1]. In recent years, the ongoing reform of English language teaching in Chinese universities has placed greater emphasis on cultivating speaking and communication skills, making it a central component of educational objectives. From traditional language teaching methods to interactive teaching approaches facilitated by modern information technology, the methods and forms of oral teaching have undergone significant changes, and research on oral proficiency has also become increasingly diverse and complex. The themes of English speaking proficiency research not only reflect the academic community's interest and trends in this field, but also, to some extent, reveal the profound influence of social and cultural changes, educational policy adjustments, and technological advancements on the practical and theoretical development of speaking proficiency improvement.

This study uses the CNKI database of core Chinese journals as its corpus and employs the LDA model to extract

and analyze topics. By interpreting these results, the research identifies key themes, trends, and driving forces in English speaking proficiency research. This paper provides a comprehensive understanding of the field's knowledge structure, offering visual insights for future research and practical guidance for English teaching and curriculum development.

2. Literature review

Oral competence refers to the ability to communicate effectively in specific contexts using spoken language, characterized by interactivity, dynamicity, and immediacy[2]. Unlike written language, spoken communication involves paralinguistic and non-linguistic features, such as intonation, pitch, facial expressions, and body language, which are crucial in actual communication[3]. While there is no consensus on the precise definition of oral competence, it can generally be categorized into two perspectives: the "cognitive process perspective", which focuses on psychological processes like language generation and monitoring[4], and the "social communicative behavior perspective", which views it as the ability to achieve communicative tasks in specific contexts, emphasizing its context-dependence and practical function[5].

As language environments diversify, oral competence has evolved to reflect multi-dimensional integration and dynamic development. Plough et al[6]. note that second language oral competence encompasses not only language knowledge and skills but also the interaction between language users and communicative contexts. Derwing et al[7]. further emphasize its dynamic nature, suggesting a shift from the traditional static view to one that incorporates developmental and interactive aspects. Wang and Zhou[8] propose a multi-dimensional model of oral competence, focusing on sub-skills, cognitive processes, and external factors, providing a more comprehensive framework for its measurement.

3. Data sources and research design

3.1. Data sources

The data for this study was sourced from the China National Knowledge Infrastructure (CNKI) database, using the keyword "oral competence" in a thematic search. To ensure academic authority and representativeness, only articles from "CSSCI" journals were selected. The search, conducted on July 29, 2025, initially yielded 285 articles. After excluding those unrelated to English oral competence, a final dataset of 146 valid articles was obtained, with the earliest Published in 1999.

3.2. Research design

This study used the titles, abstracts, and keywords of selected articles as the initial corpus. Data preprocessing was conducted using the Jieba segmentation tool to tokenize the corpus and remove stop words, ensuring model input accuracy. Multiple topic modeling methods were tested and manually reviewed for consistency. After several iterations, the LDA model was selected to identify research hotspots and academic trends in the field of "English oral competence".

The study used "topic coherence" as the primary evaluation criterion to determine the optimal number of topics. The modeling process involved two main steps: (1) calculating coherence for the initial corpus, determining the optimal number of topics, and categorizing the documents by topic, including the keywords and their weights; (2) reviewing the generated model to analyze the rationale for each document's assigned category, summarizing the main research directions, and naming the categories based on representative papers. This approach aims to identify research hotspots and provide theoretical support for future studies.

4. Research results and analysis

4.1. Topic modeling results

The study applied the LDA topic modeling technique using the Gensim library in Python to identify thematic structures in

the corpus. After testing different topic numbers and comparing their coherence scores, the optimal number of topics was determined to be three. The LDA model was then retrained, and the top 10 keywords with the highest weights for each topic were extracted, as shown in **Table 1**.

Table 1. LDA Topic Modeling Results

Topic Number	Top 10 Keywords and Weights
Topic 0	0.019*“students” + 0.017*“research” + 0.012*“in” + 0.009*“conducting” + 0.009*“English” + 0.009*“speaking” + 0.009*“English speaking” + 0.009*“teaching” + 0.008*“improvement” + 0.007*“analysis”
Topic 1	0.023*“research” + 0.012*“students” + 0.011*“speaking” + 0.008*“English” + 0.008*“speaking ability” + 0.007*“improvement” + 0.007*“language” + 0.007*“in” + 0.006*“speaking test” + 0.006*“training”
Topic 2	0.022*“research” + 0.011*“in” + 0.009*“analysis” + 0.008*“speaking” + 0.008*“speaking ability” + 0.007*“students” + 0.007*“conducting” + 0.006*“language” + 0.005*“learners” + 0.005*“based”

The LDA model effectively identified and separated the underlying themes within the corpus, with clear distinctions between topics and minimal overlap. This suggests that the model successfully captured the main themes in English speaking proficiency research. Based on the keyword distribution and sentence analysis, the three topics are named: Topic 0, “English Speaking Proficiency Pedagogy”, Topic 1, “English Speaking Proficiency Testing and Assessment”, and Topic 2, “English Speaking Proficiency Learning and Language Analysis”. Each of these themes is further discussed in the following sections.

4.2. Topic analysis

4.2.1. English Oral Competence Pedagogy

English oral competence has long been one of the central focus in foreign language education. As an interactive skill, improving oral competence requires not only mastery of language knowledge but also effective teaching strategies and practical opportunities. The design of teaching activities and classroom organization plays a crucial role in helping students apply their knowledge in real communicative contexts and enhancing their language application abilities. Therefore, pedagogy research focuses on improving language skills through scientifically designed teaching, enabling students to confidently use English in everyday communication.

The analysis of keywords shows that “teaching models” refers to the strategies and methods employed in English oral instruction. Teaching design, classroom interaction, and instructional strategies directly affect students’ oral competence. For example, Cao et al.[9] highlight the advantage of human-computer collaborative dialogue models in reducing psychological barriers to communication, providing evidence for adaptive teaching models based on subject intelligence. Huang[10] emphasizes the importance of diagnostic assessments to help teachers and students identify weaknesses in oral expression, allowing for more precise adjustments to teaching goals and alignment with evaluation standards. “Student differences” also plays a key role in English oral instruction. Variations in language proficiency, learning backgrounds, and motivations require teachers to consider individualized needs and tailor teaching strategies accordingly. Wang and Zhang[11] suggest that the significance of oral competence dimensions, such as task management and interactive listening, varies with the student’s technical context and proficiency level, with specific strategies required for problem-solving and interaction at higher levels.

Research underscores that teaching design, classroom interaction, and personalized instruction are essential for improving oral competence. Furthermore, recognizing individual differences and fostering communication skills are vital to enhancing teaching effectiveness. Future research should explore the integration of AI and personalized teaching methods to offer more scientific and comprehensive support for developing students’ English oral competence.

4.2.2. English Oral Competence Testing and Assessment

Oral competence testing and assessment are essential for operationalizing language constructs. With advancements in education and technology, English oral assessments have evolved from traditional paper-based tests to more diversified, technology-driven formats. Effective oral assessment not only measures language abilities but also promotes self-reflection and improves learning strategies.

Oral competence testing involves assessing not only fluency and accuracy but also communicative strategies, intonation, adaptability, and pragmatic competence [12]. Traditional tests often focus on grammar and pronunciation, but a deeper understanding of oral communication reveals that assessments should also consider how students respond in real communication. Key components of oral competence testing include fluency, coherence, communicative effectiveness, cultural adaptability, and interactional abilities[13].

The introduction of the "China English Proficiency Scale" (CSE) by the Ministry of Education in 2018 standardized oral testing, with a nine-level proficiency structure covering listening, speaking, reading, writing, and translation. This scale provides a framework for national language proficiency assessment. Research based on the CSE has grown, with scholars like Jie et al.[14] designing speaking tasks and scoring standards based on the CSE speaking scale, while Qi et al.[15] developed an entrance test for non-English major freshmen at a western university. Jie and Jin[16] proposed a theoretical framework for typical oral activities and language feature scales.

Today's oral assessment should not only measure language knowledge but also account for the diversity of communicative contexts and evaluate students' actual communicative competence. Future research should focus on optimizing assessment frameworks, particularly for cross-cultural communication competence and contextual adaptability. As technology advances, the integration of speech recognition and AI presents new opportunities for oral assessment, though ensuring accuracy and comprehensiveness remains a challenge.

4.2.3. English Oral Competence Learning and Language Analysis

Oral language is more immediate and interactive than written language, presenting learners with challenges in both applying language knowledge and adjusting strategies to cope with changing contexts. Improving oral competence requires mastering basic language elements such as grammar, vocabulary, and pronunciation, as well as understanding the strategies and interactions involved in oral communication. Key issues in English oral competence research include how to express ideas effectively and respond flexibly to listeners' feedback in complex contexts.

Research on "language analysis" focuses on learners' language output, emphasizing not only grammar and vocabulary but also fluency and adaptability in communication. Gao and Liu[17] analyzed discourse cohesion and coherence in spoken language, investigating features that distinguish college students' speaking levels through both quantitative and qualitative methods. Yu and Zhang[18] used complex dynamic systems theory to examine the development trajectories of oral fluency in 262 non-English major students in China, analyzing the temporal relationships between fluency subsystems. Wang and Tan[19] found a positive correlation between students' ability to vary pitch when reading words and when reading sentences, highlighting the complex relationships among oral competence dimensions. Learning English oral competence requires continuous practice to improve language sensitivity and adaptability. In discussion-type dialogues, students must not only express ideas accurately but also respond flexibly to others' questions or rebuttals. Cultivating this adaptability involves mastering language structure and demonstrating strategic language use in complex contexts. Effective oral learning depends not only on accumulating language knowledge but also on students' ability to adjust language output strategies based on communicative needs.

Future research should explore language strategies and adaptability in complex communicative contexts, especially with the integration of artificial intelligence in oral assessment, to further enhance the evaluation and development of oral competence.

5. Conclusion

This paper uses LDA topic modeling to analyze literature on "English oral competence" from core domestic journals, identifying key research hotspots, topic intensity, and evolution. The analysis reveals three main themes: English oral competence pedagogy, English oral competence testing and assessment, and English oral competence learning and language analysis.

The findings show that pedagogy research focuses on teaching methods, teacher roles, and personalized instruction; testing and assessment research centers on evaluation standards, measurement tool innovation, and cross-cultural communication competence; while learning and language analysis research examines learners' language output strategies, fluency, and adaptability. These themes reflect the academic community's multi-dimensional view of oral competence and highlight the challenges and opportunities in teaching practice.

Future research should explore optimizing oral assessment frameworks, particularly in cross-cultural communication and contextual adaptability. With advancements in AI and machine learning, AI-driven research and teaching on oral competence may become a focal point. Additionally, personalized instruction and addressing learner differences are expected to play an increasingly important role in improving oral learning and teaching effectiveness.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Luoma S, 2004, Assessing Speaking [M]. Cambridge University Press.

[2] Halliday MAK, 1985, Spoken and Written Language [M]. Geelong: Deakin University Press.

[3] Yang HZ, Zhu ZC, Fang XJ, 2011, Factor Analysis of English Speaking Ability Descriptors and Their Scaling: An Empirical Study. Modern Foreign Languages, 34(02), 151–161+219.

[4] Bygate M, 1987. Speaking [M]. Oxford: Oxford University Press.

[5] Jie W, Jin Y, 2017, A Stylistic Analysis of Oral Competence Descriptors: A Study Based on the China English Proficiency Scale. Foreign Language World, (02), 20–28.

[6] Plough I, Banerjee J, Iwashita N, 2018, Interactional Competence: Genie Out of the Bottle. Language Testing, 35(3), 427–445.

[7] Derwing TM, Munro MJ, Thomson RI, 2022, The Routledge Handbook of Second Language Acquisition and Speaking (1st ed.) [M]. London: Routledge.

[8] Wang TT, Zhou DD, 2024, A Study on the Concept of Second Language Oral Competence Based on Bibliometric Analysis. Journal of Xi'an International Studies University, 32(04), 49–55.

[9] Cao XM, Ye XL, Luo JT, et al., 2025, Research on the Psychological Barriers of Human-machine Collaborative Learning Supported by Agents—A Comparative Analysis of Multimodal Data based on an Oral English Communication Experiment. Modern Educational Technology, 35(04), 102–109.

[10] Huang Y, 2023, On the Teaching Design and Practice of Production-Oriented Approach Motivated by Diagnostic Assessment: A Lesson Study on the Unit of Story Retelling. Foreign Language Education in China, 6(04), 11–18+93.

[11] Wang H, Zhang L, 2024, The Construction of a Framework for Describing Online Oral Interaction Competence: Insights from the College English Test (CET-4 and CET-6) Speaking Exams. Foreign Language World, (05), 55–63.

[12] Dai YF, Zhang SJ, 2018, A Study on the Relationship Between Pragmatic Competence and Pragmatic Transfer in Chinese English Learners' Speaking. Foreign Language World, (01), 18–27.

[13] Wang H, Zhen FC, 2023, Online Oral Assessment of Interactional Competence: Frontiers and Trends. Modern Foreign

Languages, 46(06), 864–874.

[14] Jie W, Shao SH, Jin Y, 2024, The Application of the CSE Speaking Scale in College English Oral Learning. Foreign Language World, (06), 11–19.

[15] Qi LY, Di AY, Liu XG, et al., 2024, An Empirical Research on the Development and Validation of a CSE-based Speaking Test. Journal of China Examinations, (05), 79–89.

[16] Jie W, Jin Y, 2022, A Study on the Typical Oral Activities and Language Features Scale of the China English Proficiency Scale (CSE). Foreign Language World, (06), 2–10.

[17] Gao M, Liu Y, 2023, An Analysis of the Distinguishing Features of Discourse Performance in an Online Diagnostic Speaking Test. Journal of China Examinations, (03), 60–67.

[18] Yu HJ, Zhang GL, 2024, Dynamic Development of Oral Fluency of Chinese EFL Learners. Foreign Languages and Their Teaching, (03), 42–55+147.

[19] Wang ZJ, Tan ZG, 2023, Bottleneck and Breakthrough in the Development of Phonetic Competence of Chinese Students in College Spoken English Learning—An Experimental Study Based on the Improvement of Pitch Prominence. Technology Enhanced Foreign Language Education, (01), 76–80+115.