# An Exploration on Integration of Corpus-based Language Pedagogy into Senior High School English Curriculum

## Li Dandan[1] , Nooreen bt Noordin[2], Lilliati Ismail[3]

[1,2,3] Faculty of Educational Studies, Universiti Putra Malaysia, 43400, Serdang, Selangor, Malaysia
Corresponding Author[2]

**Abstracts**

In the globalized era, the proficiency in English language has become a pivotal skill, especially in non-native English-speaking countries like China. This study discusses the significance and methods of integrating corpus-based language pedagogy (CBLP) into senior high school English course in China. The significance of this research emanates from the increasing demand for practical, authentic, and contextually rich English language education that not only adheres to grammatical and lexical accuracy but also aligns with real-world language usage. This article is divided into seven parts. Firstly, the Corpus-based language pedagogy and Corpora, as well as Teaching methods and teaching content of Chinese senior high School English curriculum will be introduced in the literature review. Secondly, how to integrate the concept of CBLP into senior high school English curriculum will be explored. Thirdly, practical cases will be provided in the practice of integrating CBLP into senior high school English curriculum, and then evaluate the practical effects. The challenge for implementing corpus-based language pedagogy and more research directions will be summarized at last. Through the research and exploration of this article, the author hopes to provide educators with more systematic and practical teaching methods and guidance, and further improve the teaching quality of senior high school English curriculum and the comprehensive qualities of student. For future research, it is proposed to explore the longitudinal impacts of CBLP on students' overall academic performance and English proficiency, investigate the scalability of this approach in varied educational contexts, and develop a comprehensive model that can seamlessly integrate CBLP into the existing educational frameworks without necessitating substantial alterations to the current curricula.

**Keywords:** Corpus-based Language Pedagogy (CBLP), Senior High School English Curriculum, Significance and Methods, Practical Cases, Challenges and Future Prospect

## Introduction

The field of language education has undergone transformative shifts in response to the evolving demands of English language proficiency in the 21st century. As societies become increasingly interconnected and globalized, the ability to communicate effectively in English has become a prerequisite for academic success and participation in the global workforce. In China, senior high school English courses play a pivotal role in equipping students with the language skills and competencies needed to thrive in this dynamic landscape.

This paper embarks on an exploratory journey into the integration of corpus-based language pedagogy (CBLP) within the realm of senior high school English education. Corpus-based language pedagogy is a modern approach grounded in linguistic data analysis, offers a promising avenue for enhancing students' language learning experiences and outcomes. Corpus-based language pedagogy (CBLP) is similar to Shulman's pedagogical content knowledge (1987), meaning the ability to integrate corpus technology into classroom language pedagogy to facilitate language teaching. Corpus technology is intimately linked with corpus linguistics and traditionally pertains to the technological tools used in corpus linguistics for the identification of language patterns (Yoon, 2005). In our discussion of corpus technology, we place particular emphasis on its pragmatic and pedagogical applications, defining it as the utilization and implementation of technology associated with corpus linguistics and corpora for the purposes of language learning and teaching. While corpus technology encompasses a wide array of technological types and applications, our focus centers on common tools and functions, such as concordance searches, keyword analysis, frequency assessments, and the study of collocations, all of which are embedded in popular corpus websites like the Corpus of Contemporary American English (COCA), Lex-tutor, Sketch Engine, and SKELL.

In this era of information abundance, where the internet serves as an endless repository of linguistic resources, the integration of corpus-based language pedagogy holds the potential to empower students with a deeper understanding of English language structures, vocabulary, and cultural nuances. By utilizing language corpora, educators can foster critical thinking skills, cultivate linguistic accuracy, and promote cultural sensitivity among students. The corpus-based language pedagogy is perceived as revolutionary because of the authenticity and richness of its corpus-based data and its potential to promote learners as researchers who play an active role in their learning (Mukherjee, 2006). The use of corpora can encourage autonomous discovery and personalized learning (Boulton & Cobb, 2017). In a meta-analysis, Boulton and Cobb found that overall, the effect size of corpora use on learners' language improvement was large (with Cohen's d ranging from 0.95 to 1.50). What's more, corpus-based language pedagogy can help students improve various language skills, including vocabulary (Ackerley, 2017; Ma & Mei, 2021), collocations (Fang et al., 2021; Vyatkina, 2017), grammar(Lin & Lee, 2015), writing (Crosthwaite, 2020; Poole, 2016) and so on.

This former result is very encouraging because it shows the effectiveness of corpora to language learning. However, research on the practical application of corpora in the classroom has been limited (Callies, 2019; Chambers, 2019), and the conversion of theoretical findings into practical teaching designs has been relatively low. While there is a substantial amount of classroom-based corpus research, most of these studies focus on objective descriptions and analyses of existing data, lacking comprehensive discussions on corpus-based teaching models. Additionally, a majority of research targets universities, with limited focus on high school English teaching and practical teaching designs. Therefore, we need more empirical research to prove the positive impact of integrating corpora into senior high school English

teaching and to explore specific integration methods and practical cases. Such research can not only help educators better understand the actual effects of integrating corpora into senior high school English teaching, but also provide useful references and guidance for future teaching practices. Therefore, it is necessary to conduct in-depth research and exploration on the current situation of integrating corpora into senior high school English course.

**Literature Review**

This section mainly introduces the definition and history of corpus-based language pedagogy and corpora, and teaching methods and teaching content of Chinese senior high School English Courses

**Corpus-based language pedagogy and Corpora**

Corpus-based language pedagogy is an innovative approach to language teaching and learning that has gained prominence in recent years. It is founded on the principles of using linguistic corpora, which are vast collections of authentic language samples, as central resources for language instruction. This approach revolutionizes language education by grounding it in empirical data, offering learners a more practical, contextually rich, and up to date language learning experience. The development of corpus-based language pedagogy can be traced back to the late 20th century when advances in computer technology made the creation and analysis of linguistic corpora more accessible. Early pioneers, such as John Sinclair (1991), recognized the potential of corpora in language teaching. Over the years, corpus linguistics has evolved, with the integration of technology, making it a cornerstone of modern language education

Sinclair(1991) defines a corpus as a collection of language excerpts selected and arranged according to specific linguistic standards, serving as a sample of a language. Yang Huizhong (2002) views a corpus as an information repository consisting of extensive language usage data, designed specifically for linguistic research, analysis, and description. In the rapidly evolving modern society driven by computer network technology and information technology, corpora primarily refer to largescale electronic text databases obtained through scientific sampling and processing.

Since the 1980s, corpora have been widely used in all areas of language studies and have greatly advanced our understanding of language use and patterns. Linguists and researchers soon discovered the advantages of using corpora in language teaching. It is widely believed that language learners are capable of becoming 'researchers' or 'detectors' (Johns, 1991) when they engage in language learning from observing a large quantity of authentic corpus data, encouraging their discovery skills and inductive learning and enhancing their learning autonomy (Boulton & Cobb, 2017; Lewandowska, 2014). Empirical research generally supports the effectiveness of using corpora for language development. For example, the study by Johns et al. (2008) documented how 22 Chinese learners improved their English reading skills by using corpus data. Further, Boulton (2009) examined how 132 French university students made progress in learning English vocabulary after consulting corpus data. In the study by Charles (2011), 49 English as a Second Language (ESL) students self-perceived that they had increased their awareness and deepened their understanding of discourse features with the help of corpora. In addition, an experimental study by Bardovi Harlig et al. (2015) demonstrated that corpora facilitated the learning of pragmatics for 37 ESL students. Here, the experimental groups outperformed the control group by providing more appropriate oral responses with the help of corpus-informed input.

**Teaching methods and teaching content of Chinese senior high School English curriculum**

In China, the teaching methods and curriculum of high school English courses are meticulously crafted to foster students' comprehensive mastery of the English language, encompassing the four fundamental language skills: listening, speaking, reading, and writing. Within the realm of pedagogical approaches, a rich tapestry of strategies is deployed, including communicative methodologies, task-based learning paradigms, language-based gaming, flipped classroom dynamics, project-oriented pedagogy, and the integration of authentic learning materials.

As an important course for students to learn language and develop an international perspective, the teaching content is meticulously designed to encompass listening, speaking, reading, writing, grammar, vocabulary, cultural and societal dimensions, examination readiness, autonomous learning competencies, and literature. Listening exercises and activities are carefully crafted to enable students to grasp diverse accents, dialects, and speech patterns. Speaking proficiency is honed through structured discussions, debates, speeches, and roleplaying exercises. Reading comprehension spans a spectrum of textual forms, encompassing short stories, poetry, articles, and novels. Writing proficiency is cultivated through the composition of essays, the cultivation of creative writing skills, and the mastery of formal report writing. Grammar instruction is seamlessly woven into all language skills, with a distinct emphasis on applying grammatical rules appropriately within context. Vocabulary acquisition encompasses the acquisition of new words, idiomatic expressions, and the nuances of phrasal verbs. Courses in culture and civilization provide students with a comprehensive understanding of the culture, history, and societal dynamics of English-speaking countries. Examination readiness programs prepare students for standardized assessments such as the CET, as well as internationally recognized English proficiency examinations like the TOEFL and IELTS. Autonomous learning skills are imparted to students, equipping them with effective self-learning strategies, which include the proficient use of dictionaries, online resources, and language learning applications. Advanced level courses may delve into the realm of English literature, focusing on the critical analysis and interpretation of novels, drama, and poetry.

In summation, high school English curricula are meticulously designed to equip students with a robust command of the English language, instill cultural awareness, and empower them with the capacity to effectively navigate diverse linguistic contexts.

**The significance and methods of integrating corpus-based language pedagogy into senior high school English curriculum**

The development of corpus-based language pedagogy is a complex and intriguing process. In this part, the significance and methods will be introduced.

**The Significance of corpus-based language pedagogy into Senior High School English course**

The integration of corpus-based language pedagogy into English language instruction holds paramount significance as it enhances students' language learning experiences and augments the efficacy of English education. The following elucidates the importance of incorporating corpus-based language pedagogy into high school English instruction:

(1) Enriching Vocabulary and Expressions: Corpora resources provide a wealth of vocabulary, phrases, and expressions, enriching students' lexicons and enhancing their communicative prowess. Meng Chao and Ma Qinglin (2019) constructed a corpus-based legal English

vocabulary teaching model for English major students, demonstrating through empirical research that this model effectively enhances students' cognitive acquisition of legal English vocabulary.

(2) Improving Listening and Speaking: Students can enhance their listening comprehension and oral fluency by engaging with authentic dialogues and spoken language found within corpora. This is especially beneficial for preparing for oral exams and facilitating practical spoken communication. Yang Linwei and Li Ting (2015) conducted action research on multimedia news corpus self-directed learning, collecting learner feedback through questionnaires. The analysis revealed that corpora can significantly assist in listening comprehension learning.

(3) Supporting Reading and Writing: Students can utilize corpora to access relevant literature, reference materials, and sample articles, aiding in their reading comprehension and writing tasks. This proves particularly advantageous for academic writing and research projects. Xu Manfei and He An ping (2016) integrated corpus-based keyword exploration into deep reading in English major courses, demonstrating that data-driven English reading instruction enhances students' diverse reading skills and deepens their reading comprehension.

(4) Diversity and Cultural Understanding: Corpora encompass materials from various genres, contexts, and cultural backgrounds, facilitating students' understanding of how English is employed in different cultural settings. This nurtures cross-cultural understanding and respect for cultural diversity. Huang Siyun (2022) conducted a comparative analysis of the translation styles of two English translations of "The Scream" within the framework of corpus translation studies and utilizing corpus software tools. The study found that although there were stylistic differences between the two translations, both were well-received by domestic and international readers. This deepened the understanding of the profound humanistic thought within Chinese culture through the works of Lu Xun, fostering cross-cultural understanding and communication.

(5) Autonomous Learning and Pedagogical Innovation: Corpora empower students to independently seek information, engage in self-directed learning, and conduct research. This fosters students' information retrieval and critical thinking abilities. Additionally, teachers can innovate their pedagogical approaches by harnessing corpus resources to design more engaging courses and activities, igniting students' passion for learning. Gao Shuna (2016) addressed the shortcomings of the English teaching model in the automotive profession in China, proposing a "corpus-based" automotive English teaching model that integrates corpora into classroom instruction. While these teaching models have their unique features, they are all rooted in corpus-based exploration, emphasizing personalized and self-directed learning. They encourage students to discover language patterns and characteristics through the use of corpora, in stark contrast to traditional teacher-centered approaches, effectively stimulating students' enthusiasm for learning and teachers' enthusiasm for innovative instruction.

In summary, the integration of corpus-based language pedagogy into English course in China is paramount for enhancing students' language skills, cultural understanding, and practical application abilities. It contributes to more effective English education, better meeting the needs of students, and equipping them to tackle future academic and professional challenges with confidence.

**Exploring methods of Integrating corpus-based language pedagogy in into English teaching**

Incorporating corpus-based language pedagogy into high school English classrooms is a beneficial pedagogical strategy that can enhance students' English language skills and cultural understanding, among other benefits. Below are some methods for integrating corpora into high school English course:

(1) Select Appropriate Corpus Resources: Choose corpus resources that align with the course content and the students' proficiency levels, such as the COCA (Corpus of Contemporary American English) or BNC (British National Corpus). Ensure that the corpus contains texts and dialogues relevant to the course objectives.

(2) Introduce the Concept of Corpora: Introduce the concept of corpora in the classroom, explaining what it is and why it is essential for English learning. Help students understand that a corpus is a database containing a vast collection of authentic English usage.

(3) Real-Life Application: Utilize corpus resources for practical application, such as:

  *Vocabulary Studies*: Search for and analyze the usage and contexts of specific vocabulary words or phrases.

  *Grammar Instruction:* Compare examples of different sentence structures and grammar rules.

  *Listening Exercises:* Listen to real dialogues or speeches and answer related questions.

  *Reading Materials:* Provide authentic English reading materials, including news articles, blogs, and commentaries.

(4) Classroom Discussions: Encourage students to engage in discussions and share interesting English usage, cultural observations, or vocabulary they have discovered within the corpus.

(5) Writing Tasks: Require students to write essays, short texts, or reports based on corpus resources. This helps improve their writing skills and citation techniques.

(6) Cultural Exploration: Use corpus resources to explore how English is expressed in different cultural contexts, promoting cross-cultural understanding.

(7) Independent Learning: Teach students how to use corpora for self-directed learning and research. This may include how to find specific information, expand their vocabulary, and access academic resources.

(8) Assessment and Feedback: Create assessment tasks using corpus resources, such as crafting exam questions or essay prompts. Evaluate students' understanding and application of corpus resources through their work.

(9) Multimedia Tools: Leverage digital tools and online corpora, allowing students to access corpus resources in the classroom or at home to enhance their learning experiences.

When integrating corpus-based language pedagogy, teachers should adapt these methods flexibly based on students' proficiency levels and needs. These methods can help students gain a deeper understanding of practical English usage, enhance their language skills, and cultivate their self-directed learning and research abilities.

**Practical Cases of Integration of corpus-based language pedagogy into senior igh school English curriculum**

Based on the Chinese Su Teaching Edition high school English textbooks, here are two curriculum cases of Integration of corpus-based language pedagogy into Senior High School English course. Demonstrating how to integrate corpora into high school English classrooms to achieve improved educational outcomes:

*Case 1: Vocabulary Expansion*

Textbook Unit: "Environmental Issues"

Corpus Integration: Identify advanced vocabulary related to environmental problems from the corpus, such as "deforestation," "pollutants," and "ecosystem."

Lesson Design:

Lesson Objectives: Students will learn to expand their vocabulary, especially advanced vocabulary related to environmental issues, using the corpus.

Activity 1: Vocabulary Discovery

(1) The teacher introduces the day's lesson objectives, explaining that students will use the corpus to explore advanced vocabulary related to environmental issues.

(2) Students are grouped, with each group assigned a vocabulary word, such as "biodiversity."

(3) Students use the corpus to find different usages and example sentences for their assigned vocabulary and record them.

Activity 2: Vocabulary Sharing

(1) Each group of students shares the usages and example sentences they have found for their assigned vocabulary, explaining the word's meanings.

(2) The teacher leads a discussion to help students understand the vocabulary's context and cultural background.

(3) The teacher provides additional corpus-based example sentences to expand students' vocabulary knowledge.

Activity 3: Vocabulary Application

(1) Students complete an in-class task where they write a short piece describing an environmental issue, such as climate change or pollution, using the advanced vocabulary they've learned.

(2) Students review each other's pieces and provide feedback to ensure the correct application of vocabulary from the corpus.

*Case 2: Grammar Application*

Textbook Unit: "Narrative Writing"

Corpus Integration: Extract authentic narrative texts from the corpus, highlighting sentence structures, verb tenses, and narrative techniques used by native speakers.

Lesson Design:

Lesson Objectives: Students will learn to analyze and apply sentence structures and verb tenses extracted from narrative texts in the corpus.

Activity 1: Narrative Analysis

(1) The teacher presents an authentic narrative text from the corpus, emphasizing sentence structures, verb tenses, and narrative techniques.

(2) Students collectively analyze the text, discussing its features, particularly the time expressions used in narration.

Activity 2: Narrative Imitation

(1) Students are grouped, with each group receiving a different narrative text from the corpus.

(2) Students mimic the sentence structures and verb tenses found in their assigned text, creating their own narrative paragraphs.

(3) Each group shares their narrative paragraphs and discusses their choices and discoveries.

Activity 3: Narrative Writing

(1) Students write a short essay or story applying the sentence structures and verb tenses they've learned from the corpus analysis and imitation.

(2) Students exchange their work and provide feedback, emphasizing the correct application of language patterns from the corpus.

Through these detailed lesson designs, students gain a deeper understanding of and practical skills in applying vocabulary and grammar knowledge extracted from the corpus. This helps enhance their English language abilities and boosts their confidence in handling various language tasks.

*Case 3: Listening Practice*

Textbook Unit: "Travel and Tourism"

Corpus Integration: Incorporate audio clips from travel documentaries and interviews with native speakers discussing their travel experiences.

Lesson Design:

Lesson Objectives: Students will improve their listening skills by listening to authentic travel-related conversations from the corpus.

Activity 1: Authentic Listening

(1) The teacher plays authentic travel conversation clips from the corpus, such as dialogues from travel documentaries or tourist interviews.

(2) Students listen solely for comprehension without receiving the text in advance.

(3) Students discuss key information and impressions from what they heard.

Activity 2: Comprehension Tasks

(1) Students are grouped, with each group receiving a different travel conversation clip from the corpus and tasked with answering related questions.

(2) Students exchange answers, comparing their comprehension and responses.

(3) The teacher provides the conversation text for students to check and improve their answers.

Activity 3: Travel Discussion

(1) Students engage in group discussions about the content they heard in the conversations, sharing insights about travel and culture.

(2) Each group reports their discussion results, highlighting the information and insights gained from the corpus.

*Case 4: Speaking Skills*

Textbook Unit: "Social Issues and Debates"

Corpus Integration: Use the corpus to gather authentic arguments and counterarguments related to social issues like climate change, technology, or education.

Lesson Design:

Lesson Objectives: Students will practice their oral communication skills by discussing social issues and using arguments from the corpus.

Activity 1: Argument Gathering

(1) The teacher provides authentic arguments from the corpus related to social issues, including supporting and opposing viewpoints.

(2) Students independently read these arguments and prepare speaking materials.

Activity 2: Classroom Debate

(1) Students are grouped, with each group representing different viewpoints on social issues.

(2) Students debate their selected topics, using arguments from the corpus to support their positions.

(3) After the debate, students observe and comment on the language expressions used during the debate.

Activity 3: Reflection and Feedback

(1) Students review the debate, self-assess their oral expression and debate skills.

(2) The teacher provides feedback, emphasizing the effective use of arguments from the corpus and suggesting improvements in oral expression.

Through these classroom designs, students will enhance their listening comprehension and oral expression skills while gaining knowledge related to social issues. This will help nurture their critical thinking and debating abilities, enabling them to better address real-world social problems.

**Assessment methods and metrics for evaluating the impact of corpus-based language pedagogy on senior high school students**

Evaluating the impact of corpus-based language pedagogy on senior high school students requires a clear definition of evaluation methods and metrics. The details are as follows:

(1) Academic Performance

*Compare students' midterm and final exam scores before and after corpus-based teaching to determine if the teaching model has improved their academic performance.*

*Homework and project grades assess students' performance in classroom assignments and projects to see if there are improvements after corpus-based teaching.*

*Vocabulary and Grammar Accuracy: Use standardized tests or grammar assessments to evaluate students' vocabulary and grammar accuracy.*

(2) Language Skill Development

*Listening Skills:* Use listening tests to assess students' listening comprehension abilities, including their understanding of listening materials and their ability to answer oral questions.

*Oral Fluency:* Assess students' fluency and clarity in oral expression, which can be evaluated through oral speeches or conversation exercises.

*Reading Comprehension:* Use reading tests to examine students' understanding of various texts.

*Writing Skills:* Assess students' writing abilities, including organizational structure, grammar, and clarity of expression.

(3) Student Satisfaction and Feedback

*Student Surveys:* Conduct surveys to gauge student satisfaction with classroom teaching, course materials, and instructors, as well as their opinions on corpus-based teaching methods.

*Face-to-Face Interviews:* Conduct individual or group face-to-face interviews to gain deeper insights into students' learning experiences and feedback.

(4) Practical Application Abilities

*Real-World Projects or Tasks:* Assess whether students can apply what they have learned to real-life situations, such as solving practical problems or completing project tasks.

*Oral Communication:* Evaluate students' performance in real-life oral communication, including fluency, clarity of expression, and conversational skills.

(5) Cross-Cultural Communication and Cultural Sensitivity

*Cross-Cultural Projects:* Evaluate students' abilities in cross-cultural communication and understanding, including their respect for and understanding of different cultural backgrounds.

*Cultural Sensitivity:* Assess whether students have gained a deeper understanding and sensitivity to other cultures.

(6) Independent Learning and Research Abilities

*Independent Learning Projects:* Assess whether students possess the ability to learn and conduct research independently, including whether they can actively use resources such as corpora to expand their knowledge.

These evaluation methods and metrics outlined above will help provide a comprehensive understanding of the impact of corpus-based language pedagogy on students, covering various aspects such as academic performance, language skills, cultural sensitivity, practical application abilities, and independent learning capabilities. By utilizing these methods collectively, it becomes possible to comprehensively evaluate teaching effectiveness, further improve educational models, better meet students' needs, and enhance the quality of English education.

### Challenges and future prospects

This section summarizes the challenge for implementing corpus-based language pedagogy and provides more research directions for corpus-based language pedagogy for future studies.

### Challenges

Implementing corpus-based language pedagogy may offer numerous advantages in enhancing students' English language skills; however, it may also encounter several challenges, including:

(1) Access and Technological Barriers: Not all schools and students have reliable internet access and the necessary technological equipment. In some regions, students may face difficulties accessing online corpus tools, potentially limiting their learning opportunities.

(2) Credibility of Information: Students need to learn how to evaluate the credibility of information and resources within online corpora. Given the diverse nature of information on the internet, some content may be inaccurate or unreliable, necessitating instruction on identifying trustworthy sources.

(3) Varying Skill Levels: Students' English proficiency and academic capabilities differ. Using corpus tools may be overly complex for some students, requiring additional guidance and training, while others may adapt more easily.

(4) Teacher Training: Teachers need to become proficient in using corpus tools and integrating them effectively into their classrooms. They require time and training to master these skills for guiding students effectively.

(5) Resource Availability: Some schools or educational institutions may lack the resources to provide adequate corpus access or licenses, potentially impacting students' learning experiences, especially in resource-constrained settings.

(6) Alignment with Materials and Curriculum: Teachers need to invest time in integrating corpus resources into their curriculum to ensure alignment with course materials and learning objectives. This necessitates additional lesson planning and preparation.

(7) Student Motivation: Students may require extra motivation to actively engage in corpus-based teaching. If students do not fully appreciate the benefits of this teaching method for their English learning, they may not participate actively.

Addressing these challenges requires cooperation between educational institutions, teachers, and students. Key factors for successfully implementing corpus-based language pedagogy include providing adequate training and support, ensuring resource availability, safeguarding data privacy, and inspiring student interest and motivation. Moreover, the potential benefits of this teaching method, such as improving students' English skills and independent learning abilities, make it worthwhile to overcome these challenges.

### Future research directions for corpus-based language pedagogy

While corpus-based language pedagogy has made significant strides in the field of English education, several potential research directions warrant further investigation:

(1) Personalized Learning and Adaptive Education: Future research can explore how corpus tools can support personalized learning, offering customized educational experiences based on individual student needs and proficiency levels. This may involve developing adaptive education systems and gaining a better understanding of students' learning paths and requirements when using corpus resources.

(2) Multimodal Learning: Combining corpus resources with multimedia learning can enhance students' English skills. Future research can delve into how to integrate video, audio, images, and text to create richer and more engaging learning experiences.

(3) Interdisciplinary and Holistic Education: Researchers can study how corpus tools can be used across different subjects and fields to promote interdisciplinary and holistic education. This enables students to apply English skills in broader academic and professional contexts.

(4) Advancements in Educational Technology: As technology continues to evolve, future research can explore new educational technology tools and platforms to enhance corpus-based teaching models. This may include smarter search and analysis tools, as well as the application of virtual reality and augmented reality technologies.

(5) Multilingual and Cross-Cultural Education: Considering global trends, future research can investigate how corpus resources can support multilingual and cross-cultural education. This helps students better understand English usage in diverse cultural and linguistic contexts.

(6) Quality and Diversity of Corpus Resources: Future research can focus on the quality and diversity of corpus resources. This involves ensuring that the texts and data within corpora are of high quality and diverse to better meet educational needs.

These research directions will contribute to expanding the application areas of corpus-based language pedagogy and provide opportunities for educational innovation. Moreover, this research can drive a better understanding of how to effectively leverage corpus resources to support students' English learning and development.

### Conclusion

In conclusion, based on the main findings and viewpoints presented in this paper, we emphasize the potential and significance of corpus-based language pedagogy. This teaching method can assist students in better comprehending and utilizing the English language, ultimately enhancing their language skills. Schools and educational institutions should invest in teacher training to better utilize corpus resources and improve the quality of English education.

Corpus-based language pedagogy are widely recognized for their significant educational potential and value. This teaching method positively influences students' English learning and language development by providing real language experiences, expanding vocabulary, fostering text analysis and critical thinking, promoting independent learning and research, improving listening and speaking skills, supporting interdisciplinary education, and enhancing teaching effectiveness, among other aspects. Of particular importance, corpus-based language pedagogy can help students adapt to real-life English communication, improve their English language skills, cultivate critical thinking and independent learning abilities, and increase their academic and career opportunities. However, implementing this teaching method also poses challenges, including access and technological barriers, information credibility, varying student skill levels, teacher training, resource availability, privacy and data

security, educational equity, and student motivation. These challenges need to be better addressed in future research.

Future research directions encompass personalized learning, multimodal learning, interdisciplinary education, educational technology advancements, multilingual and cross-cultural education, and the quality and diversity of corpus resources. These avenues of research will further enhance the effectiveness and sustainability of corpus-based language pedagogy, propelling innovation and development in the field of English education.

**References**

Ackerley, K. (2017). *Effects of corpus-based instruction on phraseology in learner English*.

Bardovi-Harlig, K., Mossman, S., & Vellenga, H. E. (2015). The effect of instruction on pragmatic routines in academic discussion. *Language Teaching Research*, *19*(3), 324–350.

Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, *21*(1), 37–54.

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, *67*(2), 348–393.

Charles, M. (2011). Using hands-on concordancing to teach rhetorical functions: Evaluation and implications for EAP writing classes. *New Trends in Corpora and Language Learning*, 26–43.

Crosthwaite, P. (2020). Taking DDL online: Designing, implementing and evaluating a SPOC on data-driven learning for tertiary L2 writing. *Australian Review of Applied Linguistics*, *43*(2), 169–195.

Fang, L., Ma, Q., & Yan, J. (2021). The effectiveness of corpus-based training on collocation use in L2 writing for Chinese senior secondary school students. *Journal of China Computer-Assisted Language Learning*, *1*(1), 80–109.

Gao, S. (2016). Exploration of the Reform of Corpus "Data Driven" Automotive Professional English Teaching Model. *Journal of Southwest Normal University: Natural Science Edition*, *41*(10), 168–172.

Huang, Y. (2022). A Corpus Based Comparative Study on the Translation Styles of Two English Translations of "Shout." *Modern Linguistics*, *10*, 2124.

Johns, T. F., Hsingchin, L., & Lixun, W. (2008). Integrating corpus-based CALL programs in teaching English through children's literature. *Computer Assisted Language Learning*, *21*(5), 483–506.

Lewandowska, A. (2014). Using corpus-based classroom activities to enhance learner autonomy. *Konińskie Studia Językowe*, *2*(3), 237–255.

Lin, M. H., & Lee, J.-Y. (2015). Data-driven learning: Changing the teaching of grammar in EFL classes. *Elt Journal*, *69*(3), 264–274.

Ma, Q., & Mei, F. (2021). Review of corpus tools for vocabulary teaching and learning. *Journal of China Computer-Assisted Language Learning*, *1*(1), 177–190.

Meng, chao, & MA, Q. (2019). An Empirical Study on Legal English Vocabulary Teaching Model Based on Online Corpus. *Media in Foreign Language Instruction*.

Mukherjee, J. (2006). *Corpus linguistics and language pedagogy: The state of the art–and beyond. InS. Braun, K. Kohn, J. Mukherjee (Eds.) Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods (pp. 5–24)*.

Poole, R. (2016). A corpus-aided approach for the teaching and learning of rhetoric in an undergraduate composition course for L2 writers. *Journal of English for Academic Purposes*, *21*, 99–109.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1–23.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press, USA.

Vyatkina, N. (2017). *Data-driven learning of collocations: Learner performance, proficiency, and perceptions*.

Xu, M., & He, A. (2016). Corpus Technology Promotes English Deep Reading Teaching. 中国电化教育, *12*, 87–90.

Yang, H., & Wei, N. (2002). *Introduction to Corpus Linguistics*. Shanghai Foreign Language Education Press.

Yang, L., & Li, T. (2015). Research on English Listening Teaching Assisted by Network Multimedia News Corpus. *China Adult Education*, *16*, 178–179.

Yoon, H. (2005). *An investigation of students' experiences with corpus technology in second language academic writing*.