



# The Application Trends of Big Data Development in Linguistics Teaching

Xiao Yu

Kashi University, Princeton Kashi, 844000, China

18263860404@163.com

**Abstract.** This paper explores the application trends of big data in linguistics teaching. It analyzes how big data technology, with its characteristics of large volume, high velocity, and diverse variety, is gradually integrated into linguistics teaching, bringing about revolutionary changes in teaching models, learning experiences, and educational evaluation. By examining relevant research and practical cases, this paper aims to provide insights into the future development direction of linguistics teaching in the big data era.

**Keywords:** Linguistics; education; big data.

## 1 Introduction

In recent years, the rapid development of information technology has led to an explosion of data. Big data, with its unique features, has permeated various fields, and education is no exception. Linguistics teaching, which has long been centered around language acquisition, communication skills training, and cultural understanding, is now experiencing a profound transformation under the influence of big data<sup>[1-2]</sup>.

The traditional linguistics teaching model often adopts a one - size - fits - all approach, with limited consideration for individual differences among students. However, big data provides new opportunities to collect, analyze, and utilize vast amounts of educational data, enabling more personalized, efficient, and effective teaching methods.

## 2 The Characteristics of Big Data Applicable to Linguistics Teaching

### 2.1 Volume

In the context of linguistics teaching, a large volume of data can be generated. For example, students' daily language practice records, including written assignments, oral presentations, and online discussions, accumulate over time. Language corpora, which are large collections of texts, can also contain millions of words from different sources,

such as novels, newspapers, and academic papers. These large - scale data sources provide rich materials for in - depth language analysis.

## 2.2 Velocity

The real - time nature of big data is highly relevant to linguistics teaching<sup>[3]</sup>. With the popularity of online learning platforms, teachers can immediately access students' learning data, such as their participation in real - time online classes, the time spent on each learning task, and the frequency of seeking help. This real - time data allows teachers to adjust teaching strategies promptly to meet students' needs.

## 2.3 Variety

Data in linguistics teaching comes in various forms<sup>[4-5]</sup>. There are structured data, like test scores and students' demographic information, and unstructured data, such as audio and video recordings of students' language use, and text - based online chat records. The ability to handle and analyze this diverse data is crucial for comprehensively understanding students' language learning processes.

# 3 Application Trends of Big Data in Linguistics Teaching

## 3.1 Personalized Learning

Big data can be used to create detailed student profiles. By analyzing students' past learning performance, learning styles, and preferences (such as the types of language materials they like), teachers can gain a comprehensive understanding of each student. For example, a student who often reads online English news articles may have a preference for current affairs - related language learning materials. Based on this profile, personalized learning plans can be designed, including customized reading lists, vocabulary - building exercises, and grammar - focused tasks.

Adaptive learning systems are becoming increasingly popular in linguistics teaching. These systems use big data algorithms to adjust the difficulty level of learning content in real - time according to students' performance. For instance, if a student answers a series of grammar questions correctly, the system will present more challenging questions. Conversely, if a student struggles with a particular topic, the system will provide additional explanations and practice materials at an appropriate level. A study by [Researcher's Name] found that students using an adaptive learning system in a foreign language course showed a 20% improvement in overall language proficiency compared to those in a traditional class.

## 3.2 Intelligent Teaching Tools

Big data - enabled intelligent teaching tools can automatically grade students' written and oral assignments. For written work, these tools can analyze grammar, spelling,

vocabulary usage, and the overall structure of the text. They can provide instant feedback, highlighting errors and suggesting improvements. In the case of oral assignments, speech recognition technology combined with big data analysis can evaluate pronunciation, intonation, and fluency. Teachers at [School Name] reported that after implementing an automated grading system for English essays, they saved 50% of the time previously spent on grading, allowing them to focus more on providing in - depth feedback to students.

Intelligent tutoring systems are another application of big data in linguistics teaching. These systems can simulate one - on - one tutoring sessions. They use natural language processing and big data to understand students' questions and provide accurate and detailed answers. For example, when a student asks a question about a complex grammar point in an online learning environment, the intelligent tutoring system can draw on a vast database of language explanations and examples to provide a clear and relevant response.

### **3.3 Rich and Precise Learning Resources**

Big data analysis can recommend learning resources based on students' learning progress and needs. For example, if a student is learning a specific language function, such as making complaints in a foreign language, the system can recommend relevant dialogues, role - play scenarios, and cultural background materials. A language learning app, [App Name], uses big data to analyze users' search history, learning time, and performance on different topics. As a result, it can recommend highly relevant podcasts, videos, and reading materials, which has increased user engagement by 35% according to the app's developers.

Teachers can also use big data to create customized learning materials. By analyzing the common difficulties and interests of their students, they can select or produce texts, audio, and video materials that are more suitable for their class. For example, if a large number of students in a class have difficulty understanding a particular cultural concept in a foreign language, the teacher can create a short video with explanations and examples based on relevant data.

### **3.4 Integration of Language Research and Teaching**

Corpora, which are large collections of texts, are an important source of big data in linguistics. In teaching, corpus - based approaches can help students understand how language is used in real - life contexts. Teachers can use corpus data to show students common collocations, language patterns, and semantic preferences. For example, when teaching the use of the verb "make," a corpus - based analysis can reveal that "make a decision" and "make a mistake" are common collocations. A study in [Journal Name] showed that students who received corpus - based instruction had a better understanding of language usage and were more confident in their language production.

Big data research in linguistics, such as studies on language acquisition processes and language change, can be directly applied to teaching. For example, research on the most effective order of teaching grammar points based on learners' cognitive processes

can inform curriculum design. Teachers can then adjust their teaching sequences to better facilitate students' language learning.

### **3.5 Scientific and Comprehensive Teaching Evaluation**

Traditional teaching evaluation mainly relies on test scores, which have limitations in comprehensively reflecting students' learning processes. Big data allows for the collection of multidimensional data, including students' participation in class discussions, their performance in group projects, and their learning behavior on online platforms. For example, the number of times a student contributes to an online discussion forum, the quality of their comments, and their interaction with other students can all be recorded and analyzed.

Predictive analytics, a powerful application of big data, can predict students' future performance. By analyzing students' historical data, such as their learning progress in the first half of a semester, it can predict their likelihood of passing the final exam or achieving a certain language proficiency level. Teachers can then intervene early for students who are at risk of poor performance, providing additional support and resources.

### **3.6 Promotion of Interdisciplinary Teaching**

Linguistics teaching in the big data era is increasingly collaborating with other disciplines. For example, in the field of computational linguistics, linguists work with computer scientists to develop language processing algorithms. In the classroom, this interdisciplinary approach can be reflected in courses that combine language learning with data analysis skills. Students may learn how to analyze language data using statistical software while improving their language proficiency.

The integration of big data into linguistics teaching aims to cultivate interdisciplinary talents. These students are not only proficient in language skills but also have knowledge and skills in related fields, such as data science, psychology, and education. They are better prepared to meet the diverse needs of society, for example, in the fields of language technology development, cross - cultural communication research, and language - based data analysis.

## **4 Challenges and Solutions**

### **4.1 Data Privacy and Security**

Collecting and using students' data raises concerns about data privacy and security. To address this, educational institutions need to establish strict data protection policies. They should obtain students' consent before collecting data, and ensure that data is encrypted and stored securely. For example, [University Name] has implemented a privacy - by - design approach, where data protection is considered from the beginning of any data - related project in the educational setting.

## 4.2 Teacher Training

Teachers need to be trained to effectively use big data in teaching. Many teachers may not be familiar with data analysis tools and techniques. Professional development programs should be provided to help teachers learn how to interpret and utilize big data in their teaching. For instance, workshops on using learning analytics tools and understanding data - driven teaching strategies can be organized.

## 4.3 Data Quality

The quality of data is crucial for accurate analysis. Incomplete, inaccurate, or inconsistent data can lead to wrong conclusions. To ensure data quality, strict data collection and pre - processing procedures should be established. For example, when collecting students' online learning data, clear guidelines should be provided to students to ensure that they input data correctly. Figure 1 shows the number of outstanding students in the class after combining with big data and artificial intelligence. With the continuous deepening of big data teaching, there are more and more excellent students, and big data can effectively improve students' classroom efficiency.

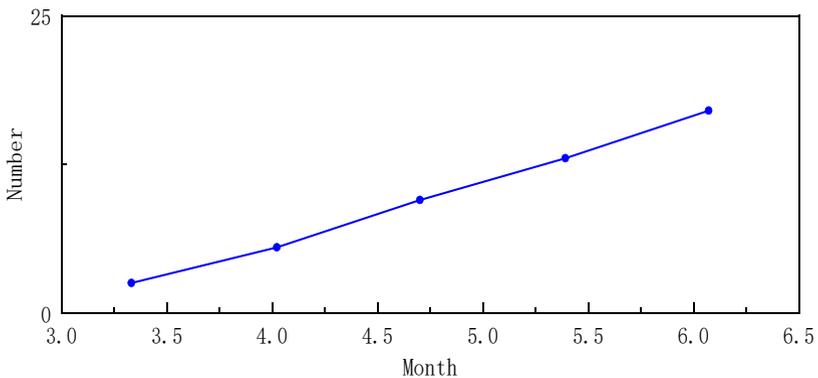


Fig. 1. Data Change Chart

## 5 Conclusion

Big data is bringing about significant changes in linguistics teaching. The trends of personalized learning, intelligent teaching tools, rich and precise learning resources, integration of language research and teaching, scientific teaching evaluation, and promotion of interdisciplinary teaching are reshaping the landscape of linguistics education. Although there are challenges such as data privacy, teacher training, and data quality to overcome, with appropriate solutions, the potential of big data in enhancing linguistics teaching is enormous. As technology continues to advance, it is expected that big data will play an even more prominent role in improving teaching effectiveness,

promoting students' language learning, and cultivating a new generation of language - proficient and interdisciplinary talents.

## References

1. Ke Q Ch, H C, Mi Q W, etc. How digitalization shapes new advantages in educational development in the construction of an educational powerhouse [J/OL]. *Modern Distance Education Research*, 1-9 [2020-01-11] <http://kns.cnki.net/kcms/detail/51.1580.G4.20250109.1741.006.html>.
2. Zhang J J. The Digitization Transformation of Psychological Health Education in Higher Vocational Education from the Perspective of New Quality Productivity: Value Implications, Realistic Challenges, and Practical Approaches [J]. *Popular Literature and Art*, 2025, (01):170-172.DOI:10.20112/j.cnki.ISSN1007-5828.2025.01.057.Author, F.: Article title. *Journal* 2(5), 99–110 (2016).
3. Zheng T T, Zeng X Y. Construction of Chinese interlanguage corpus in the era of big data [J]. *Journal of Xiamen University (Philosophy and Social Sciences Edition)*, 2016, (02): 53-63.
4. Yu L M. Finding worthwhile research questions from practice - On the practical orientation of educational linguistics [J]. *Research in Educational Linguistics*, 2024, (00): 1-7.
5. Zhang D L. Exploring the Multidisciplinary and Multimodal Nature of Educational Linguistics: A Systemic Functional Perspective [J]. *Foreign Language Guide*, 2024, 47 (01): 1-10+158.

**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.

