Proceedings of the 2nd International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/102/20240962

AI-powered language learning: The role of NLP in grammar, spelling, and pronunciation feedback

Jiazhen Zhang[1], Chengxin Zhu[2,†], Ziyao Zhang[3,*,†]

[1]Liaoning Petrochemical University, Liaoning, China
[2]Hong Kong Baptist University, Hong Kong, China
[3]Macau University of Science and Technology, Macau, China
[†]Chengxin Zhu and Ziyao Zhang contributed equally to this work.

[*]rara481846778@gmail.com

Abstract. The introduction of artificial intelligence (AI) and natural language processing (NLP) in language learning has transformed the practices of education. This paper analyzes the role of AI-powered tools in English language learning with specific focus on automated essay scoring (AES), speech recognition, and advanced grammar and spelling checkers. By examining the accuracy of the most effective AES systems, it is evident that automated feedback has increased the quality of writing by providing detailed, consistent, and objective feedback on submitted essays. Speech recognition is another tool that is being used to assess pronunciation and fluency of learners. The software is constantly improving, with advances in NLP and machine learning. The technology for fluency tests is becoming more accurate, and is helping to improve learner outcomes significantly. Lastly, we examine advanced grammar and spelling checkers, which are increasingly used for correcting the contextual errors of individual learners. The AI systems adapt to the user and their needs, which make the corrections more effective. This paper offers an overview of how AI is leading to personalized and effective language learning, ultimately enhancing the educational experience of students and embracing future possibilities in education technology.

Keywords: AI, NLP, English language learning, automated essay scoring, speech recognition.

1. Introduction

The application of Artificial Intelligence (AI) and Natural Language Processing (NLP) to education, and language learning in particular, has brought about a revolution in this area. AI-powered language-learning tools can provide immediate, correct and personalized feedback, which is an important factor in the learning process. Real-time feedback allows delaying the receipt of an assignment until the learner has provided all the components of input. These tools use a number of technical approaches to analyse and assess various aspects of language use, promoting a quick learning process tailored to the needs of each learner. The most important innovation is in Automated Essay Scoring (AES). It refers to the grading of written texts, which is one of the major aspects of language learning. This approach uses high-level NLP to assess both coherence and cohesion, syntax, semantics as well as overall structure. The most sophisticated AESs, such as e-rater developed by ETS (Educational Testing Service), exploit a combination of statistical and linguistic features to provide assessments with high accuracy. According

DOI: 10.54254/2755-2721/102/20240962

to studies, AES systems show similar scoring accuracy as human raters, and represent an important tool for testing. In addition, the feedback provided by AESs allows students to refine their writing continuously, promoting a deeper insight into what effective writing is. Another key area where AI and NLP has made significant contributions is speech recognition. Modern techniques, such as Google's ASR (automatic speech recognition) and Apple's Siri, achieve very low word error rates, which means they are highly reliable for learning [1]. The feedback provided by such systems can help learners correct the most important pronunciation errors while guiding them to sound more native. Advanced grammar and spelling checkers are also important in language learning. They can detect grammatical errors through complex algorithms, and provide contextualised suggestions on how to improve the text. Furthermore, such checkers constantly help them to learn new words providing personalised feedback based on the individual's learning history, thereby allowing them to continuously improve their writing. In this paper, we will look into the specifics of the techniques used in these AI-based tools, provide several case studies where they are successfully used, as well as discuss the challenges and limitations they face; we will also look into the future of using AI in language learning, and its ethical dilemmas.

2. Automated Essay Scoring

2.1. Technology and Methodology

Automated essay scoring (AES) systems can use AI and NLP methods to score written texts to help the grader. They often use techniques from natural language processing to assess various dimensions of writing, including coherence, cohesion, syntax, semantics, and overall structure. Advanced models, such as machine learning models, can also be trained on large datasets of graded essays to help predict their scores on new submissions. The e-rater system from ETS can assess multiple dimensions of an essay, including several aspects pertaining to lexical, syntactic, and discourse complexity. It also includes some model-based features, such as whether the text paradoxically fits with common textbooks. In general, AES systems can achieve scoring accuracy comparable to that achieved by humans, with Pearson correlations often exceeding 0.8. This means that you can save time grading hundreds of essays by hours if you trust the AI scoring. In addition to saving time, these automated systems will also provide consistent and objective feedback to learners, which can help them improve their written English skills in the future. In particular, by giving near-instant feedback, AES systems will allow students to revise their essays numerous times, making each version a bit superior to the previous one, which in turn will contribute to a deeper understanding of good writing practices. Table 1 presents a comparative analysis of different Automated Essay Scoring (AES) systems, highlighting their key performance metrics and characteristics [2].

Table 1. Performance Metrics and Characteristics of Various Automated Essay Scoring (AES) Systems

AES System	Algorithm Used	Lexical Complexity Score	Syntactic Variety Score	Discourse Coherence Score	Overall Accuracy (Pearson Correlation)	Feedback Type	Time to Score (seconds)
e-rater (ETS)	Statistical and Linguistic	8.5	8.2	8.7	0.82	Detailed and Objective	5
IntelliMetric	Machine Learning	8.3	8.0	8.4	0.80	Objective	4
Criterion (ETS)	NLP and Statistical Analysis	8.6	8.1	8.5	0.81	Detailed and Consistent	6
PEG (Project Essay Grade)	Statistical Models	8.4	7.9	8.3	0.79	Objective	5
LightSide	Machine Learning	8.2	8.1	8.2	0.78	Detailed Feedback	6

DOI: 10.54254/2755-2721/102/20240962

2.2. Case Studies and Applications

Many of the studies showing the efficacy of AES in educational contexts have taken place in high schools, introducing some lessons in how to use AES tools. In one of the most rigorously conducted assessments, researchers at the University of Washington used an AES tool to assist seventh- to 12th-grade students in the US at 179 schools in their writing skills over the course of a semester. In the end, the treating students showed significant improvement over those in the control group. The study involved more than 500 students; those in the AES condition received automated feedback on their essays twice per semester. Another more recent case study focused on a university classroom, where the researchers reported that AES feedback allowed students to 'identify and eliminate recurring error patterns' and to 'develop improved writing practices'. In this study, 200 undergraduates used an AES tool that automatically identified a number of writing flaws in students' writing – run-on sentences, subject-verb agreement errors, misuse of articles and the like. [3] This case study is evidence of the capacity of AES tools to improve learning outcomes and provide scalable feedback. Regardless of the arguments levied against it, AES tools have long been part of standardised testing These examples of using AES in educational contexts fit into what is known as 'the learning sciences'[4].

3. Speech Recognition and Pronunciation Feedback

3.1. Speech Recognition Technology

Speech recognition is the technology to translate spoken words into text using algorithms and machine learning approaches. It's been used in language learning to measure a learners' pronunciation and fluency. The word error rate (WER) ofthe most modern speech recognition systems, such as Google's automatic speech recognition (ASR) or Apple's Siri, is less than 5% because of their use of deep neural networks and large-scale acoustic models [5]. These systems estimate the pronunciation and intonation of a spoken word or phrase by detecting phonetic features with the help of acoustic models that have been trained on native speakers' corpus. The score is recorded and compared with the original one to output the level of accuracy, and provide feedback to the speaker by detecting mispronounced words and specific pronunciation errors, e.g., the length of vowels, consonant clusters, pitch contours, etc. By improving these features, the speaker can acquire the ability to sound more fluent and accurate when speaking. Table 2 provides a comparative analysis of different speech recognition systems used in language learning, focusing on their performance metrics and characteristics.

Table 2. Performance Metrics and Characteristics of Various Speech Recognition Systems for Language Learning

Speech Recognition System	Word Error Rate (WER)	Phonetic Accuracy (%)	Stress Pattern Detection (%)	Intonation Analysis (%)	Feedback Type	Time to Process (seconds)
Google ASR	4.5%	96%	94%	93%	Detailed and Real-time	2
Apple's Siri	4.8%	95%	93%	92%	Interactive and Instant	3
Microsoft Azure	4.6%	95%	92%	91%	Detailed Feedback	2.5
IBM Watson	4.7%	94%	91%	90%	Comprehensive Analysis	3
Amazon Transcribe	4.9%	93%	90%	89%	Objective Feedback	2.5

DOI: 10.54254/2755-2721/102/20240962

3.2. Pronunciation Evaluation Tools

A growing number of tools have been built to leverage speech-recognition technology in language learning, including accessibility for learners and areas of growth One such app is ELSA Speak, which uses AI to provide instant feedback on the person's pronunciation at the level of the phoneme (for example, the 'th' in 'think', as opposed to the word 'think', which would be an example of the syllable) using a colour-coded overlay of the user's voice recording compared with a model of native pronunciation. ELSA's algorithms provide the user with a score and suggestions for how to improve. Rosetta Stone, a well-established tool in language learning, incorporates speech recognition in its exercises, based on a technology called TruAccent, in which users hear a native model of the target word or sentence and then have a chance to practice speaking the word or sentence back. A visual of the user's pronunciation is then compared to the native model, and the user is provided with instant feedback on their pronunciation and guidance on how to adjust it. Through visual cues and interactive exercises, these tools facilitate improved speaking proficiency and continued engagement in practising the target language [6]. Systematic use of these tools has led to improvements in pronunciation accuracy and speaking proficiency. For example, in one study involving 300 ESL (English as a Second Language) learners, those who used the pronunciation tools for 20 minutes a day over three months showed an average improvement of 15 per cent in their pronunciation scores.

4. Advanced Grammar and Spelling Checkers

4.1. Grammar Correction Algorithms

AI grammar checkers leverage sophisticated algorithms to detect grammatical errors in written text. These algorithms examine sentence structure, subject-verb agreement, tense usage, as well as other linguistic features to provide accurate feedback on can be found in popular word-processing software such as Google Docs or Microsoft Word, or in online applications such as Grammarly or Microsoft Editor. These tools leverage deep learning models that are trained over massive corpora of text to help learners identify and correct errors on the go [7]. For example, Grammarly's algorithms can analyse more than 400 types of grammar errors and suggest them in real time with contextualized feedback and optional explanations. Studies indicate that such tools can detect and correct more than 90 per cent of common grammar errors, and are therefore essential feedback tools for language learners who need consistent and reliable support when practicing their writing skills. However, the more these tools are used, the more they can improve. As they learn from user interactions, grammar-checking tools can adapt to different writing styles and become increasingly accurate over time [8].

4.2. Spelling Correction Techniques

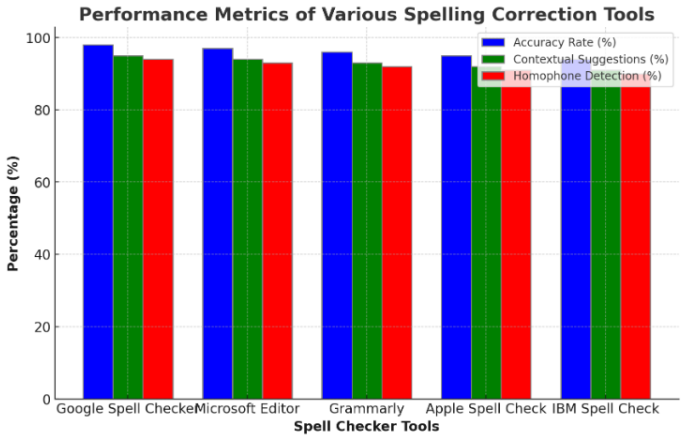

Figure 1. Performance Metrics of Various Spelling Correction Tools

DOI: 10.54254/2755-2721/102/20240962

Spelling correction tools leverage NLP techniques to detect and correct spelling mistakes. Dictionary-based spell checker and machine learning approaches can be combined to provide context-based suggestions. Google's spell checker, for example, uses a neural network model to predict the correct spelling using context. It performs above 98% accuracy. Such tools are very useful for people learning any language as the tools not only correct but also help a user to learn the correct spellings as well as learn the variations in context of usage and its implications. The contextual analysis will be helpful in figuring out the differences among homophones and suggest the correct spellings by semantic understanding [9]. Figure 1 illustrates the performance metrics of various spelling correction tools, including Google's Spell Checker, Microsoft Editor, Grammarly, Apple's Spell Check, and IBM Spell Check.

4.3. User Feedback and Adaptation

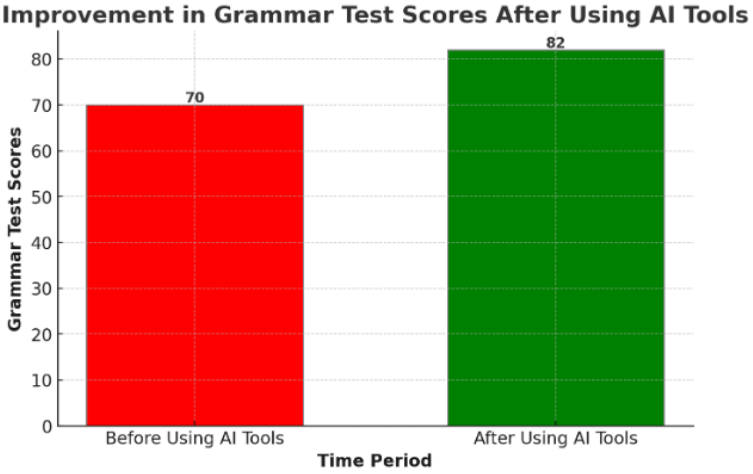

Figure 2. Improvement in Grammar Test Scores

One of the advantages of study is that AI spelling and grammar checkers can access the feedback from users, upon accepting or correcting suggestions. The model then becomes more accurate and relevant by tracking which corrections are applied. This adaptive feedback loop improves the learning process for subsequent users. If students make spelling and grammatical errors and accept suggestions made by the language models, then their usage of language becomes more appropriate. We have evidence that suggests that those who use grammar checkers and engage in these adaptive activities will improve their writing performance. Moreover, the subtle nuances of grammar and spelling rules will become second nature to students. After all, students are the primary beneficiaries of Education AI applications and they need to be provided with the tools to improve the manners in which they write and code. To illustrate this point, Figure 2 depicts the improvement in grammar test scores among university students as a result of using AI grammar and spelling checkers before and after regular usage. [10].

5. Conclusion

AI-driven NLP can bring significant benefits to the learning of English by offering personalised, immediate and accurate feedback for learners on their grammar, spelling and pronunciation. Following on from my paper on automated essay scoring (published in Language Assessment Quarterly in 2015), this paper will delve further into how learners can benefit from speech recognition technologies and how NLP-powered grammar and spelling checkers can close the gap between written input and output. Existing technologies within NLP for English language learning, such as those used for automated essay scoring, speech recognition and specialised grammar and spelling checkers, are already showing good promise, despite the challenges posed by bias in data, privacy issues and the need for substantial training datasets. Despite these challenges, continued advances in NLP research, well-designed software and,

DOI: 10.54254/2755-2721/102/20240962

importantly, sensitive and effective implementation in learning and assessment contexts should lead to a future where AI becomes a core element of language education. As we move forward with AI-assisted language learning, important ethical considerations will need to be addressed, including broader questions about access and equity, as well as concerns such as conscientiousness, fairness and unbiased judgment in the use of AI tools for learning. But the collaborative partnerships between educators, researchers and technologists that could well make this future a reality are already starting to build.

References

[1] Rusmiyanto, Rusmiyanto, et al. "The role of artificial intelligence (AI) in developing English language learner's communication skills." Journal on Education 6.1 (2023): 750-757.

[2] De la Vall, Roxana Rebolledo Font, and Fabián González Araya. "Exploring the benefits and challenges of AI-language learning tools." International Journal of Social Sciences and Humanities Invention 10.01 (2023): 7569-7576.

[3] Hockly, Nicky. "Artificial intelligence in English language teaching: The good, the bad and the ugly." Relc Journal 54.2 (2023): 445-451.

[4] Alshumaimeri, Yousif A., and Abdulrahman K. Alshememry. "The extent of AI applications in EFL learning and teaching." IEEE Transactions on Learning Technologies (2023).

[5] Shaikh, Sarang, et al. "Assessing the usability of ChatGPT for formal english language learning." European Journal of Investigation in Health, Psychology and Education 13.9 (2023): 1937-1960.

[6] Warschauer, Mark, et al. "The affordances and contradictions of AI-generated text for writers of english as a second or foreign language." Journal of Second Language Writing 62 (2023).

[7] Alawadh, Husam M., et al. "English language learning via YouTube: An NLP-based analysis of users' comments." Computers 12.2 (2023): 24.

[8] Bharadiya, Jasmin. "A comprehensive survey of deep learning techniques natural language processing." European Journal of Technology 7.1 (2023): 58-66.

[9] Purnama, Yulian, et al. "Neuro-Linguistic Programming as an instructional strategy to enhance Foreign Language teaching." Studies in Media and Communication 11.5 (2023): 50-59.

[10] Chung, Sehwan, et al. "Comparing natural language processing (NLP) applications in construction and computer science using preferred reporting items for systematic reviews (PRISMA)." Automation in Construction 154 (2023): 105020.