

# Enhancing English education with Natural Language Processing: Research and development of automated grammar checking, scoring systems, and dialogue systems

Jiazhen Zhang<sup>1</sup>, Junhui Hu<sup>2,\*</sup>

<sup>1</sup>Liaoning Petrochemical University, Liaoning, China

<sup>2</sup>Shandong University, Shandong, China

\*wellington589125@gmail.com

**Abstract.** In this paper, we investigate the scope for transformational change that these technologies offer to English education. We examine how NLP can support English education through automated grammar checking having students' text checked for grammatical errors in real time, automated scoring systems which evaluate written and verbal English against predefined criteria, and dialogue systems which interacts with learners in English to help develop their speaking and listening skills in an engaging and non-judgmental environment. We document the current status of these NLP applications, examine their educational benefits, highlight some of the challenges faced by the technologies and finally discuss the way forward for large-scale adoption of such technologies, bringing the research to a logical conclusion. We hope that this study will provide a comprehensive understanding of how NLP can be employed to transform education in English and also highlight some of the associated challenges and ethical dilemmas.

**Keywords:** Natural Language Processing, English Education, Automated Grammar Checking, Automated Scoring Systems, Dialogue Systems.

## 1. Introduction

The emergence of Natural Language Processing (NLP) has greatly facilitated advancements in various fields, including English education. NLP corresponds to a subfield of artificial intelligence focusing on the interaction between computers and human languages. NLP technologies enable computers to understand, process and analyse large amounts of natural language. In English education, NLP technologies can revolutionise the traditional methods of instruction and learning. In this paper, we will review three major applications of NLP in English education: automated grammar checking, automated scoring systems, and dialogue systems. First, we will illustrate automated grammar checking systems, which employ NLP to correct grammatical errors in text. Based on an NLP algorithm, these systems analyse the given text to detect grammatical errors and provide suggestions to improve the text quality. Automated grammar checkers exploit machine learning models trained on big corpora of text, ranging the evaluation of the text from elementary grammar rules to complex stylistic issues. Thus, learners could receive immediate feedback on the errors committed and promptly correct them. In this way, the writing skills of learners could be improved more effectively. In addition to the aforementioned

automated grammar checkers, NLP technologies can be further applied to develop automated scoring systems to evaluate written English and spoken English. Researchers have shown that people tend to evaluate writing and speaking by mixed criteria including grammar, vocabulary, coherence, pronunciation, and fluency. Based on these criteria, NLP technologies could be used to evaluate written and spoken English and provide scores upon the results of these evaluations. By utilising advanced machine learning models, automated scoring systems could process large amounts of data, including student submissions and multiple evaluations of students' works by teachers. Thus, such systems are suitable for standardised testing and cannot be assessed by human evaluators in a timely manner. Furthermore, the advances of dialogue systems, involving NLP technologies, propose new possibilities for English education. A dialogue system, also called a conversational agent or chatbot, interacts with users through natural language conversation. For English education, dialogue systems can instruct English through conversation and provide students with practice on speaking and listening via interactive conversations. The latest dialogue systems based on deep learning could engage in natural-sounding dialogue with human users and understand human-sounding input, thanks to their powerful machine learning models [1]. These systems could be integrated into diverse platforms, such as language learning software, mobile applications, and virtual reality platforms, thereby offering versatile ways for students to practise English. Overall, this paper aims to provide a concise overview of the technical foundations of these applications of NLP, current states of these researches, and the major potential of NLP technologies in changing English education. We will discuss the technological foundations, education values, challenges and limitations, as well as the future research directions of these technologies, including the integration with other emerging technologies and the ethical concerns.

## 2. Automated Grammar Checking

### 2.1. Technology Overview

Automated grammar checking systems employ NLP algorithms that analyse text essay2paragraph.com for grammatical errors, suggesting potential corrections and making recommendations to aid students in producing grammatically correct written content. The usage of machine learning models, trained on large-scale corpora of text, enable these systems to identify and correct a rich spectrum of grammatical mistakes - from simple mistakes in basic syntax rules to more complex stylistic problems. Context-aware algorithms allow systems to provide more specific and relevant feedback. They polish students' knowledge of English grammar and usage, for instance, by suggesting how to avoid using the wrong form of words like their vs there. State-of-the-art grammar checkers, for instance, tend to automatically recognise the difference between forms such as these because they are taught based on contextual cues Table 1 outlines the performance metrics of automated grammar checking systems and shows how they assist students in producing well-worded, grammatically correct essays. [2].

**Table 1.** Performance Metrics of Automated Grammar Checking Systems

| System Name  | Grammar Error Detection Rate (%) | Correction Accuracy Rate (%) | Processing Speed (words/sec) | User Satisfaction (out of 5) |
|--------------|----------------------------------|------------------------------|------------------------------|------------------------------|
| GrammarPro   | 95                               | 90                           | 200                          | 4.7                          |
| WriteRight   | 93                               | 88                           | 180                          | 4.5                          |
| SyntaxFix    | 92                               | 87                           | 190                          | 4.4                          |
| GrammarGuard | 94                               | 89                           | 195                          | 4.6                          |
| CorrectEasy  | 91                               | 85                           | 175                          | 4.3                          |

### 2.2. Educational Benefits

The most important benefit of automated grammar checking in education is the provision of immediate feedback. Grammarly, for example, corrects various grammar mistakes including subject-verb agreement, sentence structure, parallelism, punctuation, modifiers and others. Old-fashioned grammar-

checking can be time-consuming, and at times may not provide the instant response that students need in order to improve their writing on the spot. Therefore, grammar checking systems of this nature are perfect for instant error correction and revision during writing, as students can view their mistakes immediately and make adjustments as they go. Finally, grammar checking systems like Grammarly can be used across different educational levels such as primary to higher education, making them a flexible tool in the English learning ecosystem [3]. In a high school English class, for example, students who use grammar checkers can be expected to draft essays with fewer errors. As a result, teachers can provide more feedback on the content and ideas as opposed to the simple corrections. The immediate feedback loop can also be beneficial in terms of learning, as students can see the direct impact of their corrections and can eventually be more effective in retaining grammatical rules.

### 2.3. Challenges and Limitations

Even with the strides made through NLP, grammar checking systems still have many limitations. One major weakness is the inability to handle contexts or idiomatic expressions that are relatively common in natural language and idiosyncratic for algorithms to recognise. For instance, a idiom such as ‘kick the bucket’ or ‘spill the beans’ can be misinterpreted by a system whose focus is on the literal meaning of words. Additionally, the system might struggle with the diversity of English dialects and style of writing by suggesting incorrect conventions or missing errors. Overcoming these problems requires continuous research to refine the algorithms as well as to expand the linguistic datasets used for training. For example, exposing the system to a wider range of dialectal variants and stylistic contexts in training data can help address the weaknesses [4]. Additionally, users’ ongoing feedback and live testing in the real world can help to identify these weaknesses and eventually train grammar checkers to be more sophisticated and adaptive to context.

## 3. Automated Scoring Systems

### 3.1. Technology Overview

Although these systems use NLP techniques to score written and spoken English, they’re not natural language models because their automated scoring is grounded in criteria defined by teachers and instructors, and they can provide objective and consistent assessments of written and oral skills. Automated scoring systems score written and spoken assessments on a range of language features, including grammar, vocabulary, coherence and fluency, and pronunciation. By deploying machine learning models, these systems quickly process and score large number of student submissions. These systems work well for standardised testing and large-scale tasks such as grading tests on which large numbers of students are writing. For instance, automated scoring systems such as the e-rater® developed by the Educational Testing Service (ETS) to evaluate the writing of individuals taking the GRE and TOEFL can score essays within a few seconds using implicit models and score the essay in a way that’s very similar to human raters. Such models are developed and trained on large amounts of data that reflect a variety of writing samples, providing a way to assess writing on a spectrum from low- to high-level. Table 2 shows the performance indicators of the automatic scoring system [5].

**Table 2.** Performance Metrics of Automated Scoring Systems

| System Name | Scoring Consistency with Human Raters (%) | Processing Speed (submissions/sec) | Assessment Dimensions (Grammar, Vocabulary, Coherence, Pronunciation) | User Satisfaction (out of 5) |
|-------------|---|------------------------------------|---|------------------------------|
| ScoreMaster | 92  | 50                                 | Grammar, Vocabulary, Coherence, Pronunciation                         | 4.6                          |

**Table 2.** (continued).

|             |    |    |   |     |
|-------------|----|----|---|-----|
| EvalPro     | 90 | 55 | Grammar, Vocabulary,<br>Coherence                   | 4.5 |
| AssessRight | 91 | 60 | Grammar, Vocabulary,<br>Coherence,<br>Pronunciation | 4.7 |
| GradeTech   | 89 | 52 | Grammar, Vocabulary,<br>Coherence                   | 4.4 |
| RateEase    | 93 | 58 | Grammar, Vocabulary,<br>Coherence,<br>Pronunciation | 4.8 |

### 3.2. Educational Benefits

The use of automated scoring systems for English education comes with many benefits. Firstly, it eliminates lots of variations and subjectivities accrued from human grading. The impartial and consistent nature of the system ensures that everyone is treated equally, regardless of other factors affecting the examiners. Secondly, it significantly reduces the workload of educators and enable them to focus more on one-to-one teaching and tuition. For instance, a teacher can use the automated score to deduce which student needs more help and which needs a change in approach. Thirdly, the ease with which large-scale assessments can be carried out quickly allow for more accurate measurement of learning outcomes and timely decision making to improve academic programs [6]. This is particularly important for large institutions or institutions offering online learning with hundreds of thousands of students submitting work for assessment. Human grading would be inefficient and impractical.

### 3.3. Challenges and Limitations

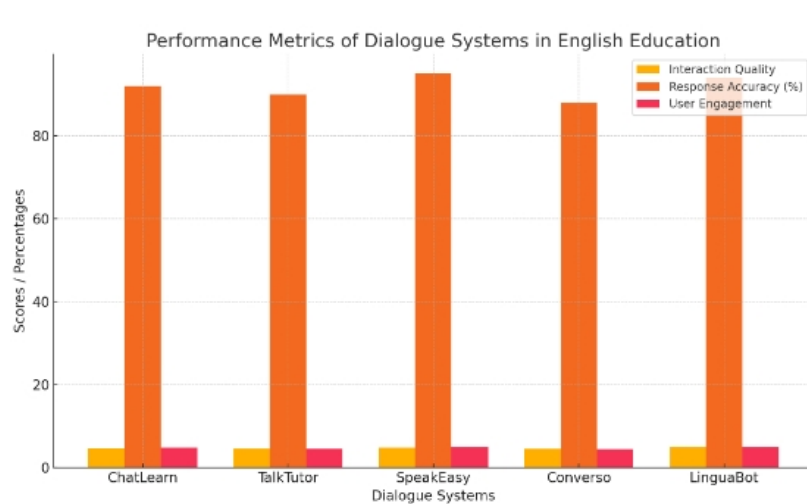
The big promise of automated scoring of essays is obvious – consistency, speed and, most importantly in our eyes, accuracy. The big pitfall is the same – consistency, speed and, most importantly in our eyes, accuracy. Consider, for example, the wide range of assignments, not only in terms of genre, but in terms of expectations for content, organisation and complexity. In short, automated scoring is much more likely to be accurate for writing assignments that focus on simple and straightforward tasks. Creative or complex writing are much harder to evaluate automatically, and automated systems might have a harder time appreciating high-level crafting of an essay, such as use of rhetorical devices or sophisticated argumentation. For example, an essay using sarcasm or irony might be graded poorly because a computer system could not detect its very subtle tone. Second, although automated systems might provide an opportunity to evaluate students' writing on a higher frequency, there is risk of over-reliance in assessment practices, and ultimately erosion of teaching of critical thinking and analysis. To mitigate these types of pitfalls, our scoring algorithms will continue to be refined, and human oversight in the grading process will always be an important element of the system.

## 4. Dialogue Systems

### 4.1. Technology Overview

NLP is widely used in dialogue systems, or conversational agents and chatbots, which simulate human-sounding conversations with users and can be effectively employed in English education to engage students in interactive dialogues to practise their speaking and listening skills. Advanced dialogue systems make use of deep learning models to understand and generate responses in natural language and can have a meaningful conversation and provide contextually appropriate responses in everyday situations. These systems can be built into a language learning platform, mobile apps and virtual reality environments and offer different memorable ways for students to practise English [7]. For instance, a dialogue system embedded into a mobile app can stimulate shopping experience, in which the student

needs to interact with a chatbot as if he/she is able to talk to a store clerk to buy things, practising corresponding real-word language skills. Figure 1 illustrates the performance metrics of various dialogue systems used in English education.[8]



**Figure 1.** Performance Metrics of Dialogue Systems in English Education

#### 4.2. Educational Benefits

The utilization of dialogue systems in English education generates several obvious benefits. First of all, these systems provide the learners with a safe and non-judgmental space to practice their language skills. Consequently, it lowers the level of anxiety associated with learning a new language. Next, dialogue systems adapt the conversation to students' level and/or learning goals. In other words, they are tailored to the needs of the language learners providing them with the relevant and challenging practice for successful learning [9]. For example, a beginner learner can practise simple, slow and short dialogues with limited vocabulary, while an advanced learner can participate in fast-paced discussions on complex topics. Moreover, language learners can practise with dialogue systems at any time, day or night. Hence, these systems boost the level of accessibility to learning resources. This particular benefit supports the continuous nature of learning by providing the learners with the opportunity to practice their language skills at any point in time. [10].

#### 5. Conclusion

The use of NLP in English education has transformed the learning experience. Automated grammar checkers, automated scoring and dialogue systems all give students immediate feedback; fair, consistent and objective feedback with consistent application of rules; and many opportunities for interactive practice with graded feedback and feedback on progress. Automated grammar checkers highlight errors and provide explanations, while at the same time prompting learners to use the grammatical rule correctly. Automated scoring systems give fair and objective assessments of written and spoken English and lessen the grading burden on teachers. Dialogue systems can provide learners with practice in spoken and written English in a safe, comfortable, personalised system that gives rapid feedback to practice opportunities. Dialogue systems can also be tailored to an individual student's level and learning goals, from total freedom in writing to semi-restricted systems with feedback on errors. In spite of these benefits, there are some challenges. Contextual use and idiomatic use of English, as well as dialectal differences, all present problems for automated systems. Similarly, there are concerns about how to keep users engaged with an intelligent system and whether assessments can be accurate if they are automated. For complex and creative writing tasks especially, there are many checks needed, too many for a human assessor to complete, even over a period of hours. Therefore, it is important that these systems are continually tested for their effectiveness, and improved upon as they are used. The future of NLP in the

education of English looks very bright. This exciting technology has many more opportunities for further development and integration with other emerging technologies such as artificial intelligence, virtual reality and augmented reality. All of these in combination would take learners beyond textbook and classroom learning environments by offering more immersive and interactive teaching and learning opportunities.

### **Contribution**

Jiazhen Zhang and Junhui Hu: Conceptualization, Methodology, Data curation, Writing- Original draft preparation, Visualization, Investigation.

### **References**

- [1] Bauer, Elisabeth, et al. "Using natural language processing to support peer - feedback in the age of artificial intelligence: a cross - disciplinary framework and a research agenda." *British Journal of Educational Technology* 54.5 (2023): 1222-1245.
- [2] Bharadiya, Jasmin. "A comprehensive survey of deep learning techniques natural language processing." *European Journal of Technology* 7.1 (2023): 58-66.
- [3] Caplar R and Kulisic P 1973 *Proc. Int. Conf. on Nuclear Physics (Munich) vol 1* (Amsterdam:North-Holland/American Elsevier) p 517
- [4] Alqahtani, Tariq, et al. "The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research." *Research in Social and Administrative Pharmacy* (2023).
- [5] Parde, Natalie. "Natural language processing." *The SAGE Handbook of Human–Machine Communication* (2023): 318.
- [6] Khurana, Diksha, et al. "Natural language processing: State of the art, current trends and challenges." *Multimedia tools and applications* 82.3 (2023): 3713-3744.
- [7] Olayiwola, Abisola, Dare Olayiwola, and Ajibola Oyedeji. "Development of an automatic grammar checker for Yorùbá word processing using Government and Binding Theory." *Expert Systems with Applications* 236 (2024): 121351.
- [8] Kumar, Ishan. *Development of Grammar Checker for Hindi Sentences*. Diss. NITJ, 2023.
- [9] Almusharraf, Norah, and Hind Alotaibi. "An error-analysis study from an EFL writing context: Human and Automated Essay Scoring Approaches." *Technology, Knowledge and Learning* 28.3 (2023): 1015-1031.
- [10] Alharbi, Wael. "AI in the foreign language classroom: A pedagogical overview of automated writing assistance tools." *Education Research International* 2023.1 (2023): 4253331.