

# Natural Language Processing in Grammar Checker

<sup>1</sup>Matthew N. O. Sadiku and <sup>2</sup>Janet O. Sadiku,

<sup>1</sup>Roy G. Perry College of Engineering, Prairie View A&M University, Prairie View, TX, USA

<sup>2</sup>Juliana King University, Houston, TX, USA

**Abstract:** Grammar is a set of systematic rules wherein language can be composed or expressed. Grammar correction is crucial for essential documents such as research, business articles, and formal papers. The advent of natural language processing (NLP) revolutionized this landscape, enabling grammar checkers to move beyond simple pattern matching to a deeper understanding of text. Natural language processing (NLP) is a subfield of artificial intelligence that focuses on the interaction between computers and human languages. One of its most pervasive and practical applications is the development of grammar checkers. This paper explores the transformative applications of NLP in modern grammar checking.

**Keywords:** *Natural Language Processing, NLP, Computational Linguistics, Grammar Checker, Spelling Checker*

## I. INTRODUCTION

In an age where written communication is paramount, grammar checkers have become indispensable tools for students, professionals, and casual writers alike. Early grammar checkers primarily operated on a set of predefined linguistic rules. The advent of natural language processing (NLP) revolutionized this landscape, enabling grammar checkers to move beyond simple pattern matching to a deeper understanding of text. NLP is a branch of artificial intelligence that focuses on the interaction between computers and human language. By combining computational linguistics, machine learning, and deep learning, NLP allows machines to process, analyze, understand, and generate human language. This means moving from merely flagging incorrect spellings to comprehending sentence structure, semantic meaning, and even the intended tone of a piece of writing [1].

The act of reading and interpreting language is known as natural language understanding (NLU). By removing metadata from text, such as concepts, entities, keywords, emotions, relations, and semantic roles, NLU enables computers to comprehend and interpret human language. The act of writing or generating language is known as natural language generation (NLG). NLG serves as a translator, turning the computerized data into a representation in natural language. It produces non-linguistic outputs from natural language inputs [2]. These components enable NLP machine learning to analyze and understand text, making it valuable for various applications.

## II. FUNDAMENTALS OF NLP

Natural language processing is a subfield of artificial intelligence that empowers computers to understand, interpret, and generate human language. It is a technique where machine can become more human and thereby making human to communicate with the machine easily. NLP seeks to make software intelligent enough to process a natural language as humans. For example, imagine a machine that takes instructions by voice.

NLP analysis generally consists of the following three levels [3]:

- *Syntax*, the study of sentence structure. Syntax deals with the formation of a sentence from individual words. Syntax alone suggests the proper interpretation of “Jimmy loves Lucy.”
- *Semantics*, the study of context-independent meaning. This derives the meaning of a sentence based on the meanings of the words/phrases. For example, semantics determines whether the word “bank” refers to a river bank or to a financial institution.
- *Pragmatics*, the study of context-dependent meaning. Pragmatics deals with how meaning changes in the presence of a specific context and how the contexts affect the meaning of the sentences. This level is concerned with the purposeful use of language in situations.

As a foundational pillar of modern artificial intelligence, NLP encompasses a wide array of tasks, including speech recognition, text classification, natural language understanding (NLU), and natural language generation (NLG). NLP encompasses a wide range of tasks, such as information retrieval (IR), named entity recognition (NER), relation extraction, text classification, topic modeling, semantic textual similarity, machine translation, and question answering (QA). Figure 1 shows how NLP transforms raw acoustic data into meaningful interactions [4], while Figure 2 shows different components of NLP [5].

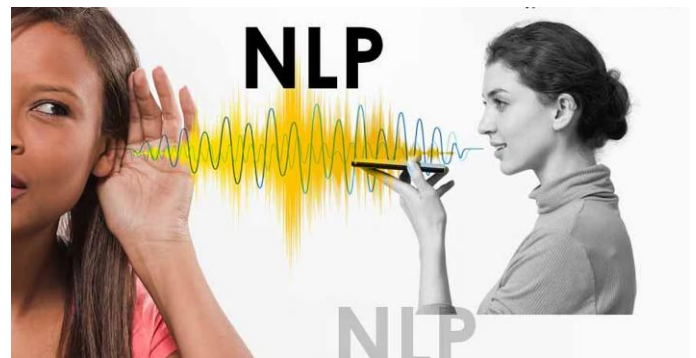


Figure 1: A representation of NLP [4].

Recently, large language models (LLMs) have shown their ability in learning universal language representations, text understanding and generation. LLMs refer to a model with a large number of parameters, vast training data, and substantial compute, enabling it to capture complex language patterns. In LLM-based NLP, pre-processing is followed by prompt engineering, which guides LLMs to produce outputs that align with extraction requirements during inference without altering the model's parameters. Models like GPT are pushing the boundaries of language understanding, enabling nuanced and context-aware applications. The GPT (Generative Pretrained Transformer) is a large-scale language model developed by OpenAI that consists of multiple layers of transformer blocks,

each with a self-attention mechanism and a forward neural network [6]. GPT-based systems can summarize complex reports or generate creative content like essays, making them versatile in both academic and professional environments. ChatGPT uses NLP techniques to understand prompts. When you enter a prompt, the chatbot comprehends it and provides relevant replies.

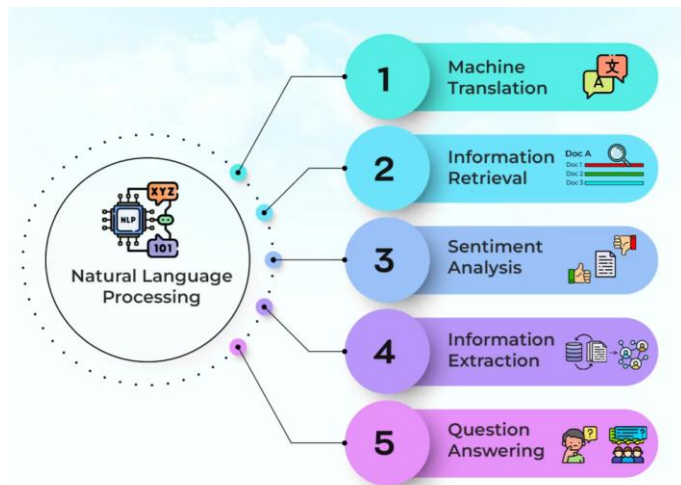


Figure 2: Different components of NLP [5].

### III. WHAT IS GRAMMAR CHECKER?

Grammar is a set of systematic rules wherein language can be composed or expressed. Grammar correction is crucial for essential documents such as research, business articles, and formal papers. There have been created grammar correction online, which has helped most students create academic writing. The use of grammar checking tools is to improve the sentence structure used in a document. The ability to write a paper with minimal to no grammatical errors is a skill requiring experience. On the other hand, language in linguistics refers to sounds or gestures that people create that is often accompanied by grammar rules wherein it has five main domains: Pragmatics, Syntax (sentence structure), Morphology (structure of words), Semantics (meaning or definition), and Phonology (the sounds of language). Each of these components should be addressed to achieve accurate error recognition and grammar improvement [7].

Grammar checkers reduce the cognitive load on writers, allowing them to allocate more mental resources to higher-order concerns such as content development, argumentation, and structural coherence. This can be particularly beneficial for second language learners or individuals struggling with basic grammar and mechanics. The widespread adoption of grammar checkers has had a profound impact on writing practices and education. These tools can significantly improve the immediate quality of written texts. Proponents argue that they serve as valuable learning aids, providing instant feedback and helping students identify and understand their common errors. Grammar checkers can foster incidental learning and empower students to become more independent editors of their work [1].

Traditional grammar checkers relied on “hard-coded” rules—static sets of instructions that flagged specific patterns as errors. The technological evolution of grammar checkers mirrors the advancements in computational linguistics. What began as rudimentary spell-checkers in early word processors has evolved into sophisticated writing companions capable of understanding nuance, tone, and complex syntactic structures. At the heart of this transformation lies natural language processing (NLP), a subfield of artificial intelligence that

bridges the gap between human communication and machine understanding.

### IV. NLP IN GRAMMAR CHECKER

The earliest grammar checkers, emerging in the late 1970s and early 1980s, were primarily rule-based. Rule-based systems are limited by the foresight of their creators; they can only detect errors that have been explicitly defined as such in their codebase. They lacked an understanding of sentence structure and were often criticized for high false-positive rates. As computational power increased, the field shifted toward statistical natural language processing (NLP). Machines can comprehend, analyze, manipulate, and interpret human languages because of NLP technology. Instead of relying solely on hardcoded rules, these systems used probabilistic models to determine the likelihood of a word sequence. The mid-2010s marked the beginning of the deep learning era in NLP. The current state-of-the-art in grammar checking is dominated by the transformer architecture, which allows the model to weigh the importance of different words in a sentence regardless of their distance from one another. Modern tools like Grammarly and Language Tool leverage these architectures to provide not just corrections, but stylistic suggestions and tone adjustments.

The most significant recent advancement in NLP for grammar checking comes from transformer models. Architectures like BERT (Bidirectional Encoder Representations from Transformers), GPT (Generative Pre-trained Transformer), and T5 (text-to-text transfer transformer) have revolutionized how language is processed. Modern tools do more than fix broken sentences; they improve the overall quality of writing [1].

### V. BENEFITS

NLP-driven grammar checkers serve as a vital bridge for non-native speakers. The greatest benefit of NLP is not just the elimination of errors, but the empowerment of writers to communicate their ideas with greater precision, confidence, and impact. Other benefits of NLP in grammar check include the following [1]:

- *Efficiency:* In a fast-paced digital environment, the speed of communication is paramount. NLP allows for real-time processing, providing instantaneous feedback as a user types. This immediate reinforcement helps prevent the accumulation of errors and reduces the time spent on final revisions.
- *Scalability:* The scalability of NLP models means they can handle vast quantities of text across various platforms—from email clients to social media and professional IDEs. This ubiquity ensures a consistent standard of quality across all forms of digital communication, saving countless hours of manual proofreading for professionals and students alike.

### VI. CHALLENGES

Grammar checkers present both opportunities and challenges. Despite significant advancements, NLP in grammar checking still faces hurdles. There are challenges of ambiguity, context, data scarcity, data quality, latency, and ethical bias. Language is dynamic; slang, technical jargon, and evolving dialects often challenge models trained on formal datasets. There is an ongoing concern regarding algorithmic bias, where models may inadvertently enforce specific cultural or regional norms over others. Beyond the linguistic and data-centric hurdles, there are significant technical challenges in deploying these

systems at scale. Other challenges of NLP in grammar check include the following [1]:

- **Ambiguity:** The primary challenges stem from the inherent ambiguity of human language, the subtle nuances of context, and the technical limitations of current computational models. A significant hurdle for NLP models is distinguishing between correct and incorrect usage when a sentence is syntactically valid but semantically or pragmatically flawed.
- **Complexity:** The most fundamental challenge in grammar checking is the sheer complexity of natural language. Unlike mathematical equations, language is governed by rules that are often fluid, context-dependent, and subject to cultural shifts.
- **Ethical Concerns:** The widespread use of grammar checkers introduces ethical concerns. These systems often enforce a specific “standard” version of a language, typically based on Western, middle-class norms. This can inadvertently marginalize dialects (such as African American Vernacular English) or regional variations (such as Indian or Singaporean English) by labeling them as “incorrect.”
- **Data Quality:** If the training data does not reflect real-world usage, the resulting model may be highly accurate in a laboratory setting but fail when faced with the “noisy” input of a diverse user base.
- **Data Limitations:** The performance of any NLP system is heavily dependent on the quality and diversity of its training data. In the realm of grammatical error correction, data-related challenges are particularly pronounced.
- **Style:** A grammar checker must distinguish between a formal academic paper, a creative poem, and a casual social media post. Non-standard syntax in poetry or the use of slang in dialogue are often flagged as “errors” by systems that lack a deep understanding of intent and style.
- **Digital Divide:** Most available datasets are concentrated in high-resource languages like English, leaving low-resource languages with significantly less effective tools. This creates a digital divide where speakers of minoritized languages do not have access to the same level of writing assistance as those of dominant languages.
- **Latency:** Modern LLM-based grammar checkers are computationally expensive. Running a multi-billion parameter model for every keystroke in a web browser or word processor introduces significant latency. Developers must often compromise between model complexity (accuracy) and speed (user experience).
- **Over-correction:** Grammar checkers operate on a spectrum between precision and recall. A system with high precision avoids flagging correct text as wrong but may miss subtle errors (under-correction). Conversely, a system with high recall catches more mistakes but risks annoying the user with false positives (over-correction).
- **Over-reliance:** Critics express concerns about potential over-reliance, fearing that students might become dependent on these tools and neglect to develop their own grammatical understanding and editing skills. While grammar checkers can highlight errors, they may not always explain the underlying

grammatical rules, potentially hindering deeper learning.

## VII. FUTURE OF NLP GRAMMAR CHECKER

The evolution of NLP in grammar checking has been remarkable, moving from simple dictionary lookups to context-aware neural architectures. The future of grammar checkers is intrinsically linked to the ongoing advancements in artificial intelligence, especially in NLP. The field of NLP is continuously evolving, and with it, the capabilities of grammar checkers. The future of grammar checkers lies in personalization. Future systems will likely adapt to an individual's unique writing style, providing suggestions that align with their personal voice while maintaining grammatical integrity.

Future advancements are likely to include even more sophisticated contextual understanding, real-time feedback that adapts to a user's unique writing style, and seamless integration into various writing platforms. As NLP continues to advance—incorporating even deeper levels of reasoning and cross-lingual understanding—the boundary between human and machine editing will continue to blur. As these tools continue to shape how we communicate, the goal must be to empower human expression rather than merely enforcing mechanical conformity [1].

## CONCLUSION

Grammar checkers have evolved from simple utility programs to powerful AI-driven writing companions. Modern grammar checkers employ a variety of NLP techniques to achieve their impressive accuracy and contextual awareness. They have become indispensable tools for global communication, fostering clarity and precision in the digital age. Natural language processing has transformed grammar checkers from simple spell-check tools into powerful writing aids. Modern grammar checkers, powered by deep learning and large language models (LLMs), assist millions of users daily in producing clear and professional text.

An AI grammar checker is a tool that automatically detects and fixes grammar, spelling, punctuation, clarity, and style issues in your writing. If you want error-free, polished, and professional writing nowadays, you need the right AI grammar checker tools on your side. In fact, using a grammar AI checker is no longer a dispensable luxury; it is an indispensable necessity. Modern grammar AI checker tools use natural language processing (NLP) and machine learning to understand your sentence structure and offer human-like corrections. Some tools help with not only grammar but also spelling, punctuation, sentence structure, readability, and more. More information about NLP grammar checker can be found in [8-10] and the following related journals:

- *Natural Language Processing Journal*
- *Journal of Emerging Technologies and Innovative Research*

## References

- [1] <https://manus.im>
- [2] P. I. A. Devi and A. A. Saleema, “Spell checking and correcting techniques in NLP,” *Futuristic Trends in Computing Technologies and Data Sciences*, vol. 3, 2023, pp. 154-161.
- [3] J. Hirschberg, B. W. Ballard, and D. Hindle, “Natural language processing,” *AT&T Technical Journal*, Jan./Feb. 1988, vol. 67, no. 1, 1988.

- [4] A. Jain, "What is the role of NLP in voice assistants?" August 2024, [https://www.analyticsinsight.net/nlp/what-is-the-role-of-nlp-in-voice-assistants#google\\_vignette](https://www.analyticsinsight.net/nlp/what-is-the-role-of-nlp-in-voice-assistants#google_vignette)
- [5] "How Google uses NLP to improve SERPs, featured Snippets & UX," <https://digitalguider.com/blog/what-is-google-nlp/>
- [6] X. Jiang et al., "Applications of natural language processing and large language models in materials discovery," *NPJ Computational Materials*, vol. 11, no.79, 2025.
- [7] P. M. A. Cano et al., "Natural language processing of grammar checker tools for academic writing: A systematic literature review," *Journal of Electrical Systems*, vol. 20, no. 7, 2024, pp. 1151-1156.
- [8] S. Vajjala et al., *Practical Natural Language Processing*. O'Reilly Media, 2020.
- [9] H. Lane and M. Dyshel, *Natural Language Processing in Action*. Manning Publications, 2nd edition, 2025.
- [10] R. C. Dougherty, *Natural Language Computing: An English Generative Grammar in Prolog*. Psychology Press, 1994.

#### ABOUT THE AUTHORS

**Matthew N.O. Sadiku** is a professor emeritus in the Department of Electrical and Computer Engineering at Prairie View A&M University, Prairie View, Texas. He is the author of several books and papers. His areas of research interest include computational electromagnetics, computer networks, engineering education, and marriage counseling. He is a Life Fellow of IEEE and a fellow of NAE.

**Janet O. Sadiku** holds bachelor degree in Nursing Science in 1980 at the University of Ife, now known as Obafemi Awolowo University, Nigeria and doctoral degree from Juliana King University, Houston, TX in December 2023. She has worked as a nurse, educator, and church minister in Nigeria, United Kingdom, Canada, and United States. She is a co-author of some papers and books.