

Paperpal

AI-driven solutions to revolutionise global research publishing

For 20 years, Cactus Communications have been providing end-to-end publication support to academic writers across the globe. Paperpal is their latest innovation in this field, providing instant AI-driven feedback on language and grammar issues, as well as technical checks more specific to the needs of a research manuscript. Research Outreach spoke with Charlotte Baptista, Product Head at Paperpal, about academic language editing, the ways that AI improves on this, and the issues it has yet to overcome.

Academic English editing services are crucial in enabling researchers from around the world to overcome language barriers and make their research accessible to all. Since 2002, Cactus Communications have provided these services and are now embracing innovation with Paperpal. This tool provides tailored language edits and technical checks for academic manuscripts, driven by artificial intelligence (AI), machine learning (ML), and Natural Language Processing (NLP) to help manuscripts reach publication. Paperpal aims to build on this technology, adding additional language editing features to the tool and improving on the AI that underpins it.

What led you to develop Paperpal and how does it work?

For several years, automation, AI, and ML have been buzzwords across the

industry. Editing was starting to take the same route that translation did; AI translation was becoming increasingly commonplace among researchers who speak English as a second language, and interest and confidence in AI editing was also growing. Cactus Communications already had two decades of experience offering manuscript preparation and publication support services to researchers across the globe. We were in a position to leverage our deep understanding of research writing and journal publication and combine it with a strong technology team of editorial subject matter experts, data scientists, and AI and NLP developers to create smart solutions.

Paperpal was born out of these efforts and created with the intention to transform the manuscript preparation experience for researchers. Our mission is to empower authors to report their research through a high-quality manuscript that can guarantee improved publication outcomes. With Paperpal, authors can get relevant academic-focused writing suggestions on their manuscript or draft in minutes. Unlike other tools in the market, Paperpal's AI is enriched with a layer of intelligence that helps it understand the anatomy and conventions of research writing. This ensures the language suggestions are suited to scientific writing tone and style. Paperpal is also equipped with a large suite of technical

checks that can help researchers validate and guard against non-compliance issues that could lead to desk rejection.

What are the most common errors that researchers make when writing research papers?

Cactus Communications has previously published our findings on common mistakes found in research manuscripts authored by Japanese and Chinese researchers. The error analysis was performed on 1,000-word samples taken from 100 papers each written by Japanese and Chinese authors. We find that researchers from these countries predominantly tend to make grammar errors, with article usage, prepositions, and subject-verb agreement being the top three types of mistakes. After grammar, clarity of expression was the second most common focus area overall; non-native speakers of English find it challenging to describe the science effectively in English, due to which some sentences seem illogical and confusing. Figures 1 and 2 are adapted from the analysis focusing on Chinese and Japanese authors; a similar trend has been observed in authors across Southeast-Asian countries.



With educational equity being of upmost importance, having an accessible tool to navigate academic research writing has become increasingly vital.

Paperpal was created with the intention of transforming the manuscript preparation experience for researchers.

How does instant app-based feedback help researchers with writing academic manuscripts?

Paperpal comes with several advantages. Firstly, all our solutions provide instant language feedback as opposed to the turnaround of a proofreading service, and it also beats waiting for colleagues to return with feedback! Secondly, features like contextual synonyms and academic writing tips which are built into our real-time solutions, Paperpal for Word and Paperpal for Web, can help assist authors who are not confident in writing in English. Finally, the technical checks report from the Paperpal for Manuscript service can act as an early-warning system to help researchers catch mistakes that result in desk rejection; both language suggestions and technical check flags are automatically applied to a document using tracked changes, giving authors an immediate sense of the level of intervention required before the manuscript is ready for journal submission.

How are advances in AI leading the revolution in editing?

Before breakthroughs in AI for grammar error correction in the mid-2010s, most early solutions were built using programmatic rules, and there were many limitations. Rules tend to fail due to the variation we see in English sentence structure. For example, there are a large number of string patterns in which a subject-verb error could manifest in a sentence, and it's difficult to detect and correct this error across every scenario.

There were also difficulties scaling the corrections to catch errors beyond basic grammar into clarity-focused rephrases, the way human copyeditors work. The earlier solutions were also unable to detect expressions and sentences that did not sound natural enough. Lastly, all error corrections worked without any understanding of the larger sentence or paragraph-level context. All of this

resulted in low accuracy and a low volume of tool suggestions. I believe this is why tools like Grammarly tend to produce conservative and error-prone correction for academic writing; their grammar checker appears to have been developed largely before the advances in machine learning.

In comparison, AI solutions like Paperpal that have learned from corrections made by expert academic editing professionals using large training datasets across domains have demonstrated significant editing coverage and a lower error rate. The neural network on which AI models are based functions similarly to the human brain, forming automatic patterns and connections. This results in AI editing outputs that often match, or at least come very close to, human achievement.

What do you think researchers, institutes, and publishers should be looking for when considering incorporating



Charlotte Baptista,
Product Head
at Paperpal.

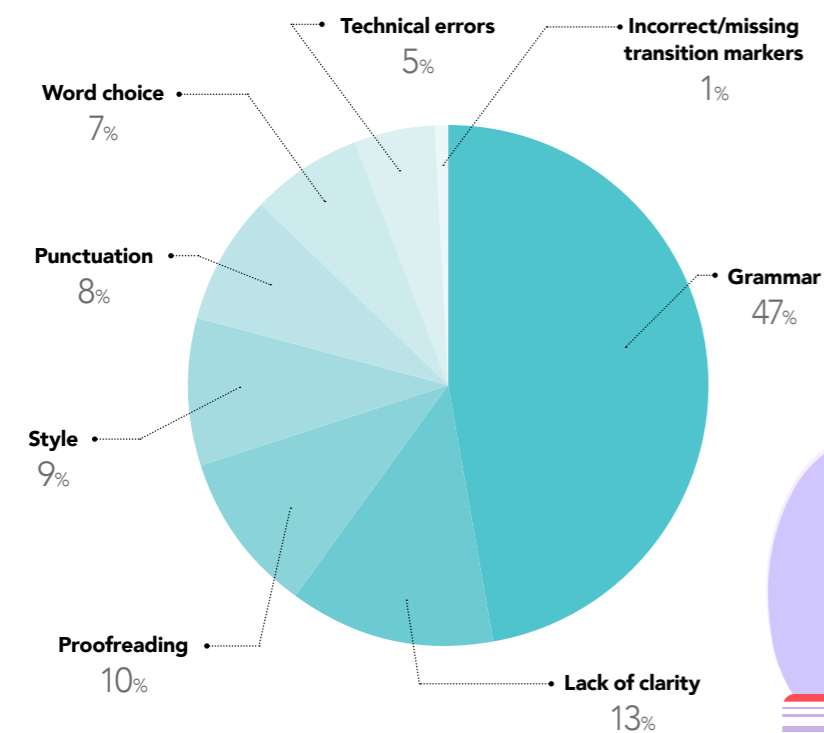


Figure 1. Common mistakes found in research manuscripts authored by Chinese researchers.
Photo Credit: Modified from Editage (2019) 100 common mistakes in research manuscript preparation, Tsinghua University Press. cactusglobal.com/press/editage-and-tsinghua-university-press-release-a-book-on-common-mistakes-in-manuscript-writing-for-chinese-researchers

Before breakthroughs in AI, automated grammar checkers worked without any understanding of the larger sentence- or paragraph-level context.

AI-assisted technology to aid research publications?

The first concern for many researchers and other end users is that the output of AI editing may not be accurate, particularly when handling technical entities in research papers and preserving meaning. In their search for the most accurate and useful tool, it is important to ensure that the solution they choose has been tailored to academic writing. Paperpal's technology team has risen to the challenge here.

Some researchers may object to automated tools as they feel they do not offer the breadth of functionality required for their research writing needs. For instance, if an author writes in a language other than English they

may use machine translation, and then go on to process the English-translated text further with grammar checkers, dictionaries, paraphraser, and other tools. Hence, such AI tools should cater to a range of needs in the writing process beyond grammar and spelling. Paperpal gives authors access not only to language correction, but also contextual academic synonyms validated from the published literature as well as AI translation from more than 25 languages to English.

Finally, researchers incorporating AI-based solutions should guard against unrealistically high expectations of the outputs. While AI technology will keep improving over time, carrying out a careful review of the outputs is a

good safety net for now. Researchers with lower English proficiency who are unable to validate the outputs of AI tools themselves, should always have the final manuscript undergo a close inspection before the draft is submitted to the journal or put to some other use.

What still needs to be overcome in machine learning to improve AI-assisted writing tools?

There are three key issues that still limit the output of AI-based tools. These elements feature heavily in the Paperpal development roadmap.

Consistency: Machine learning results sometimes tend to be hit or miss; as a result, the performance can vary across sections, specific paper types, and domains. The training dataset and machine learning model needs to be able to address these limitations.

Accuracy: Some noise suppression is always necessary to improve the final result; in other words, the AI solution should be enriched with adequate intelligence to help it understand how to approach finer details in an academic paper. This can help save hours spent in reviewing and manually rejecting incorrect suggestions when it comes

to terminology, numbers, equations, units, and so on.

Meaning: This is the toughest limitation yet. Unless AI solutions develop a strong capability to understand the context within a paper and in the real world, the performance of AI models that help with sentence prediction and text generation, particularly the new GPT models, will be found lacking.

Please tell us about your long-term goals at Paperpal.

We have prepared an exciting roadmap for the coming year to offer a more integrated experience. We understand that researchers often switch between multiple tabs to access more than one writing tool. We have recently made additional features such as AI translation available to make the writing process more efficient. Over time, our goal is to keep adding useful features using cutting-edge technology, to make the product accessible to more researchers across platforms, and to make it easy for them to collaborate while writing. We are confident that as the product evolves, Paperpal can continue to have a positive impact on their publication journey and help them achieve career success.

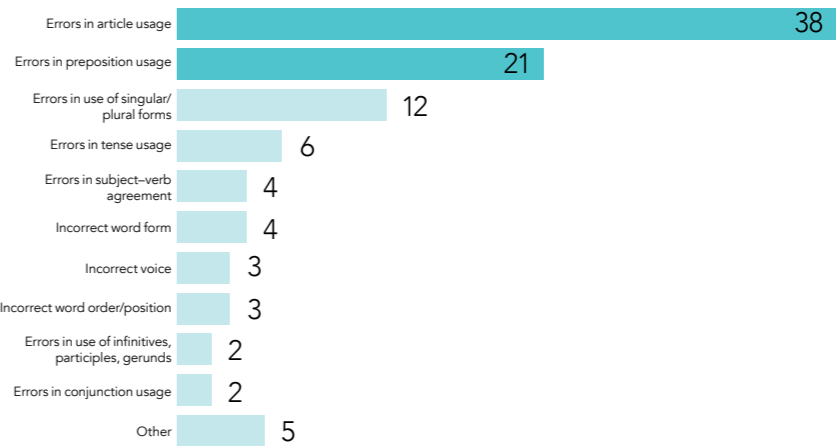
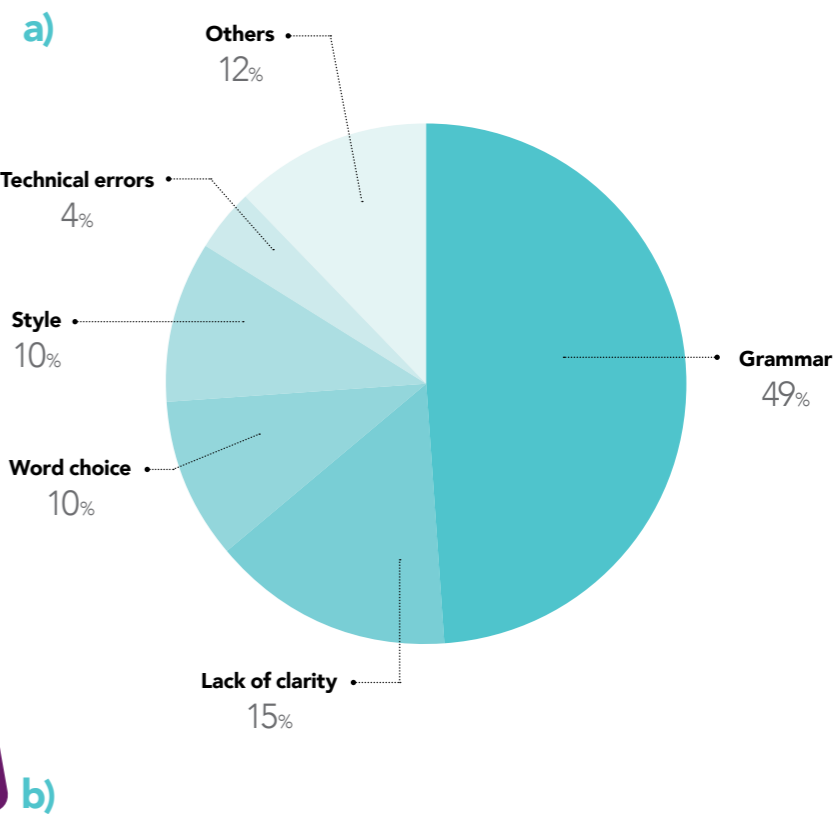


Figure 2. a) Common mistakes found in research manuscripts authored by Japanese researchers. b) Percentage of different types of grammar errors.

Photo Credit: Modified from Editage (2017) Common mistakes in research writing across subject areas: Advanced advice from an English editing company, The Japan Times. <https://www.editage.jp/blog/the-editing-companys-guide-to-english-paper-mistakes-book-is-now-available/>

Charlotte Baptista spent a decade of her career empowering large, global copyediting teams to create publish-worthy research papers. During this time, Baptista experienced first-hand the challenges that researchers face at different points in their publication journey. Eventually, she moved to the tech space, where she now leverages these insights and builds products to solve the problems that research writers encounter from idea to draft and through to publication. Baptista is passionate about automation for success and speed, and enjoys very much the process of bringing AI-enabled efficiencies to the publishing sector.

www.linkedin.com/in/charlottebaptista
W: paperpal.com

