

**Mariia Shchedrina***

PhD Student, Lecturer

Institute of Higher Education of the National Academy of Educational Sciences of Ukraine

01014, 9 Bastionna Str, Kyiv, Ukraine

Nanyang Polytechnic

569830, 180 Ang Mo Kio Ave., Singapore, Singapore

<https://orcid.org/0000-0002-0214-2024>**Yue Man Judy Yip**

Senior Lecturer

Nanyang Polytechnic

569830, 180 Ang Mo Kio Ave., Singapore, Singapore

<https://orcid.org/0009-0002-3377-6609>

Evaluating ChatGPT in Japanese language education and essay assessment

Abstract. The relevance of the study stems from the growing need to introduce artificial intelligence into the educational process, particularly in teaching non-Latin languages, such as Japanese. The combination of optical character recognition and AI-based analysis opens up new opportunities to improve the efficiency of writing assessment, the quality of feedback, and the individualisation of learning. This study aimed to investigate the use of ChatGPT to improve the effectiveness of Japanese language teaching through the automated checking of student essays. Methodologically, educators used the OCR function of ChatGPT to digitise handwritten student essays, enabling analysis of grammar, vocabulary, sentence structure, and kanji. A mixed-methods approach was adopted, combining quantitative error analysis with qualitative insights from instructors. The study revealed that ChatGPT effectively detected eight major error categories, with the most common being insufficient kanji usage, overuse of basic sentence structures, and a limited vocabulary range. The tool's ability to generate instant error summaries and targeted feedback significantly reduced the time educators spent on marking, ensuring consistent and individualised support for learners. However, the study also identified limitations, including occasional inaccuracies in OCR processing, especially for stylised or unclear handwriting. Despite these challenges, the findings suggested that ChatGPT can serve as a valuable complementary tool to traditional teaching methods, enhancing feedback quality, supporting differentiated instruction, and promoting learner autonomy. The findings offer practical guidance for language educators, curriculum designers, and edtech developers seeking to integrate AI into the teaching and assessment of less commonly taught languages, such as Japanese

Keywords: generative AI; language proficiency evaluation; AI in education; automated feedback; Japanese language pedagogy

INTRODUCTION

The integration of artificial intelligence (AI) into education has opened new pathways for enhancing teaching and learning across disciplines. Among the most widely

discussed tools is ChatGPT, a large language model developed by OpenAI, which has shown significant potential in supporting language acquisition through real-time

Suggested Citation:

Shchedrina, M., & Yip, Y.M.J. (2025). Evaluating ChatGPT in Japanese language education and essay assessment. *Scientia et Societis*, 4(2), 28-36. doi: 10.69587/ss/2.2025.28.

*Corresponding author



feedback, error correction, and learner engagement. While much of the discourse around AI in education has focused on English or widely taught European languages, there remains limited research on its application in the context of Japanese as a foreign language, particularly in assessing student writing and supporting educators in managing feedback.

At Nanyang Polytechnic (NYP) in Singapore, students have the opportunity to study Japanese as an optional elective. These modules attract self-motivated learners from various academic backgrounds, creating a dynamic learning environment where innovation in pedagogy is highly encouraged. However, Japanese poses unique challenges due to its complex writing systems (hiragana, katakana, and kanji), context-dependent grammar, and nuanced vocabulary. These challenges often manifest in student essays, where issues with coherence, kanji accuracy, and grammar can hinder effective communication.

Recent scholarly efforts have explored the multifaceted role of ChatGPT in education, extending beyond surface-level implementation to examine its pedagogical, ethical, and technological implications across diverse learning contexts. Researchers such as D. Baidoo-Anu & L. Owusu Ansah (2023) have highlighted the influence of ChatGPT on academic writing practices, prompting a reconsideration of assessment and originality in higher education. In parallel, Y.K. Dwivedi *et al.* (2021) examined how generative AI is transforming business education, specifically in terms of curriculum design and instructional methods. Concerns about academic integrity have been raised by T. Susnjak & T.R. McIntosh (2024), who warned of ChatGPT's potential to enable AI-assisted cheating, thereby challenging the validity of traditional assessment models.

Technical studies have also begun to evaluate ChatGPT's functional effectiveness in discipline-specific learning. The model's perceived educational utility among pre-service teachers has been explored by M.A. Pesce & D. Fernández Blanco (2024), who identified both promise and scepticism in its classroom applications. Similarly, I.S. Chaudhry *et al.* (2023) argued that the emergence of ChatGPT challenges the effectiveness of traditional performance-based evaluation systems in higher education, necessitating a fundamental re-evaluation of assessment methods, academic integrity safeguards, and institutional policies to ensure the ethical and purposeful integration of AI tools into learning.

Inclusion and accessibility have also emerged as themes of interest. G.B. Mohan *et al.* (2024) proposed that ChatGPT can support inclusive education by simplifying complex content. M. Li (2024) demonstrated that ChatGPT can reliably assist in marking written assessments, particularly subjective tasks such as reflective writing, when guided by well-crafted prompts and averaged outputs. However, its limitations in assessing objective or low-quality submissions indicate that it should function as a supporting tool rather than a sole evaluator. Ethical dimensions and institutional readiness were explored by F. Kamoun *et al.* (2024), who assessed faculty attitudes and concerns regarding the

integration of ChatGPT into academic workflows. In terms of instructional innovation, K. Woodruff *et al.* (2023) contributed to the K-12 discourse by identifying the practical barriers faced by teachers adopting the tool in early education. Regional applications have also been studied: M.T. Pham & T.X.T. Cao (2025) reviewed ChatGPT's use in blended language learning in Southeast Asia, noting both infrastructural challenges and pedagogical gains. K.T. Kotsis (2024) critically evaluated ChatGPT's accuracy in answering science-related conceptual questions, comparing its output with that of conventional educational tools.

Together, these studies reflect a growing interest in applying AI and OCR technologies to support Japanese language education, particularly in the assessment of handwritten student work. While significant progress has been made, ongoing research continues to refine these tools to better serve the evolving needs of educators and learners. The research ultimately aimed to evaluate ChatGPT's potential as a complementary tool to traditional teaching methods for enhancing student learning outcomes.

MATERIALS AND METHODS

This study employed a mixed-methods research design to examine the effectiveness of ChatGPT in evaluating handwritten Japanese essays and supporting educators through automated feedback. The research was conducted within the Japanese Optional Elective Level 3 (OE3) module at Nanyang Polytechnic, Singapore, and involved 28 student participants from diverse academic backgrounds who had opted to study Japanese as a foreign language. As part of the classroom task, students were instructed to write essays by hand on the topic 私のたんじょうび ("My Birthday"), drawing upon the vocabulary, kanji, and grammatical structures acquired during the course. Once collected, the handwritten essays were scanned and digitised using ChatGPT's OCR technology, which enabled the text to be processed in a machine-readable format suitable for AI-driven analysis. These digitised texts were then analysed using the GPT-4 version of ChatGPT through a tailored prompt that guided the model to identify common linguistic issues and generate feedback.

Specifically, the model was instructed to detect grammatical errors (e.g. misuse of particles, incorrect verb forms), lexical limitations (such as repetition of basic vocabulary and lack of variation), syntactic simplicity (e.g. overuse of elementary sentence patterns), and kanji-related issues (such as the omission or misuse of kanji characters). Based on this analysis, ChatGPT provided two forms of feedback: individual feedback tailored to each student, offering correction suggestions and learning tips, and cohort-level summaries that identified frequent errors and writing trends across the group. A qualitative content analysis was conducted on the AI-generated feedback to assess its clarity, specificity, and pedagogical relevance. A quantitative error frequency analysis was employed to categorise and count the most prevalent types of errors identified by the model. This methodological combination enabled the

study to evaluate ChatGPT's capability not only in identifying linguistic errors in handwritten Japanese texts but also in offering meaningful, formative feedback that could inform instruction and enhance student learning.

During the study, the ChatGPT generative tool (version GPT-4, developed by OpenAI) was used to analyse digitised handwritten texts. The model was accessed between April and September 2024 using a query designed to identify grammatical errors, repetition of basic vocabulary, simplified syntactic constructions, and errors in the use of kanji. The AI-generated feedback included individualised advice to students as well as generalised error patterns observed within the sample. The authors of the study were responsible for verifying the content and interpreting the results. No generative AI tools were used to create visual materials.

RESULTS AND DISCUSSION

This study examined the application of ChatGPT in analysing handwritten Japanese essays and providing individualised and group-level feedback. It contributes to the growing body of research on AI in education, offering new insights into its application in the context of Japanese as a foreign language – an area that remains underexplored. Nanyang Polytechnic, Singapore, offers four foreign languages – French, German, Japanese, and Korean – as optional elective modules. Students can choose any of these languages based on their interests, academic goals, or career plans. As these modules are not mandatory, students who enrol are often assumed to be more intrinsically motivated than those in compulsory courses. These optional electives, commonly referred to as OEs, provide students with the flexibility to explore subjects beyond their core curriculum. They also bring together learners from various schools across the polytechnic, including the School of Engineering, Applied Sciences, Information Technology, Business Management, Health and Social Sciences, and Design and Media. This diversity contributes to a rich and engaging learning environment built around a shared interest in language learning.

The OE modules are structured across four levels. Levels 1 and 2 offer two-hour weekly sessions and carry two credits, while Levels 3 and 4 involve four-hour weekly classes and are worth four credits. Since October 2022, the Centre for Innovation and Life Skills (CILS) has adopted a hybrid teaching format for these modules, with 40% of sessions conducted live online and 60% held in person. This blended approach aims to enhance flexibility, promote active engagement, and better prepare students for a future shaped by remote work and digital collaboration.

The integration of AI tools such as ChatGPT for assessing handwritten essays offers several pedagogical benefits. Firstly, the use of optical character recognition (OCR) technology ensures that handwritten submissions are accurately digitised, minimising the risk of transcription errors that could arise from manual input. This allows instructors to focus on providing meaningful feedback rather than spending time on basic transcription tasks. Secondly, the AI's ability

to generate personalised comments and suggestions for improvement ensures that students receive targeted feedback specific to their language proficiency level. This individualised approach aligns with best practices in language education, where tailored feedback is recognised as a key factor in learner development. Furthermore, the automated analysis of essays enables consistent and unbiased assessment, thereby reducing the likelihood of subjective grading discrepancies. The scalability of AI-assisted assessment also makes it feasible for educators to efficiently manage large class sizes without compromising the quality of feedback.

Despite its advantages, the use of AI tools for assessing handwritten essays also presents several challenges. One primary concern is the accuracy of OCR technology, which can struggle with poorly written or stylised handwriting, resulting in incorrect transcriptions. Additionally, while ChatGPT can provide relevant feedback, there is a risk that its suggestions may not always align with the instructor's pedagogical intentions – particularly if the AI-generated feedback is overly generalised or contextually inaccurate. Educators must, therefore, maintain oversight of the AI's feedback to ensure that it aligns with course objectives. Another significant challenge involves the ethical consideration of data privacy, as student essays uploaded to an AI system must be handled in compliance with institutional data protection policies. Lastly, there is the potential for over-reliance on AI-generated feedback, which may reduce the instructor's direct engagement with student writing and limit the development of critical teacher-student interactions in language learning.

A practical demonstration of AI-assisted text recognition and transcription in the Japanese Optional Elective Level 3 course involved a handwritten essay by Student A on the topic “私のたんじょうび” (“My Birthday”). This essay was uploaded to ChatGPT, where OCR technology accurately converted the handwritten text into a digital format. Figure 1 demonstrates that the digital transcription closely matches the original handwritten content, achieving a 99% accuracy rate. This high level of accuracy highlights the reliability of AI tools in digitising student submissions, preserving textual integrity while enabling efficient processing and feedback generation. Figure 2 presents another example of AI-assisted text recognition and transcription applied to a handwritten Japanese essay submitted by Student B, focusing on the same topic, “私のたんじょうび” (“My Birthday”). The handwritten text was digitised by ChatGPT, resulting in a digital transcription displayed on the right. Unlike the previous example with a 99% match, this transcription achieved approximately 90% accuracy, as shown in Figure 2. The highlighted sections in the digital text reveal discrepancies, such as incorrect or missing characters, which reflect the limitations of OCR technology in accurately recognising handwritten text – particularly when dealing with variations in handwriting style, character spacing, or alignment. These errors underscore the importance of maintaining instructor oversight when utilising AI tools for educational assessment, as automated

transcription can misinterpret student input, potentially compromising the quality of feedback. The reduced accuracy in this case also emphasises the need for consistent handwriting standards or preprocessing techniques to

optimise OCR performance. This example provides a critical perspective on the reliability of AI-assisted assessment, illustrating that while technology can significantly enhance assessment efficiency, it is not without its limitations.

Essay of Student A

The student chose the topic 私のたんじょうび

4. Essay
Compose an essay of one of the topics below. At least 12 sentences.

- ① 好きな人 (好きなひと)
② 私のたんじょうび (わたしのたんじょうび)

Topic: 私のたんじょうび

私はナンヤン ボリテクニクの三年生です。
ビジネスマネジメントが主の学生です。
私のたんじょうびは 3月 21日です。
それとまた 12時にオーチャード駅で会いました。
それから わたしは えきから ラーメン 店へ
いきました。
みそは あまり 大きく ありませんが、ラーメンは とても
おいしかったです。
たべた後で、カラオケに 行きました。
カラオケのへや は 大きくて、たのしかったです。
あとで、わたしは 私のたんじょうびに かんじます。
はが ケーキ を 買って、わたしは 私のたんじょうびを
いきました。
そして、かき が 私に プレゼント を あげました。
私のたんじょうび は とても たのしい、おもしろい です。

私はナンヤン ボリテクニクの三年生です。
ビジネスマネジメント学部の学生です。
私の誕生日は3月21日です。
私と友達12時にオーチャード駅で会いました。
それから、私たちは駅から4階のラーメン屋まで歩きました。
店はあまり大きくありませんが、ラーメンはとても美味しかったです。
食べた後で、カラオケに行きました。
カラオケの部屋は大きくて、楽しかったです。
あとで、私たちは私の家に帰ります。
母がケーキを買って、私たちは私の誕生日を祝いました。
そして、家族が私にプレゼントをあげました。
私の誕生日はとても楽しかった。特別な日です。

99% match

Figure 1. Essay by Student A and its transcribed version by ChatGPT

Source: ChatGPT

Essay of Student B

The student chose the topic 私のたんじょうび

4. Essay
Compose an essay of one of the topics below. At least 12 sentences.

- ① 好きな人 (好きなひと)
② 私のたんじょうび (わたしのたんじょうび)

Topic: 2

わたしのたんじょうびは七月にしています**。
かぞくともだちからのゆくふくやプレゼントがたの
しみです。
たんじょうびというべつなじょうりをたり**、がい
ゆくりします。
毎年、友だちがうちに来るなら、プレゼントをもって
きます。
それは私の大好きなほんです。
ほかのたんじょうびはかぞくといっしょに12さいです。
お母さんはわたしがよいこどもであるようにケーキを
つくります。
友だちといっしょにおいおいをします。
私のたんじょうびにはたくさんのお金が**。
それはやちがたんじょうびであり、わたしのじんせ
いにおけるたねびであります。
私はその日をさいしゅうにまちっています。
たんじょうびはかぞくや友だちからのあいじょうやサ
ポートをかんじることができるのでとてもくべつな
日です。

わたしのたんじょうびは七月にしています**。
かぞくともだちからのゆくふくやプレゼントがたの
しみです。
たんじょうびというべつなじょうりをたり**、がい
ゆくりします。
毎年、友だちがうちに来るなら、プレゼントをもって
きます。
それは私の大好きなほんです。
ほかのたんじょうびはかぞくといっしょに12さいです。
お母さんはわたしがよいこどもであるようにケーキを
つくります。
友だちといっしょにおいおいをします。
私のたんじょうびにはたくさんのお金が**。
それはやちがたんじょうびであり、わたしのじんせ
いにおけるたねびであります。
私はその日をさいしゅうにまちっています。
たんじょうびはかぞくや友だちからのあいじょうやサ
ポートをかんじることができるのでとてもくべつな
日です。

~90% match

Figure 2. Essay by Student B and its transcribed version by ChatGPT

Source: ChatGPT

Essay of Student C

The student chose the topic 私のたんじょうび

Topic: 私のたんじょうび

私のたんじょうびはじゅうがつにじゅうさんにちです。
じゅうぎょうのあとで、ともだちは私のうちにきました。
みんないっしょにひるごはんをたべて、ケーキをたべ
ました。
ケーキはとてもおいしかったです。
私はともだちからおきいおもちゃをもらいました。
そのあとで、私たちは家でてつてすていかにいきました。
あたらしいけいをして、しずかな世でおちやのあそびました。
ともだちは私にでがけをあげました。
いっしょにでがけをよんで、パンをたべました。
私のともだちはとてもしんせつとおもいました。
ごじゅうじにうちにえきました。
あのひはとてもたのしかったです。

私のたんじょうびはじゅうがつにじゅうさんにちです。
じゅうぎょうのあとで、ともだちは私のうちにきました。
みんないっしょにひるごはんをたべて、ケーキをたべ
ました。
ケーキはとてもおいしかったです。
私はともだちからおきいおもちゃをもらいました。
そのあとで、私たちははかってにちかいデパートにい
きました。
あたらしいけいをして、しずかな世でおちやの
のみました。
ともだちは私にでがけをあげました。
いっしょにでがけをよんで、パンをたべました。
私のともだちはとてもしんせつとおもいました。
ごじゅうじにうちにえきました。
あのひはとてもたのしかったです。

~99% match

Figure 3. Essay by Student C and its transcribed version by ChatGPT

Source: ChatGPT

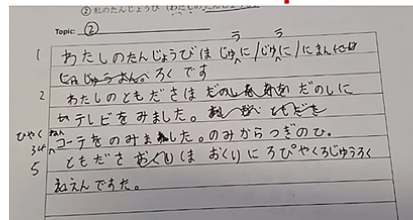
The high accuracy is evident in the minimal discrepancies between the handwritten text and the digital output presented in Figure 3, with only a minor difference highlighted in the phrase “はかつてに” (mistakenly digitised), reflecting the system’s capacity to accurately recognise and replicate student writing. The result further supports the argument that AI-driven transcription can significantly enhance assessment efficiency, reducing the manual effort required for text digitisation while preserving the authenticity of student submissions.

The text recognition applied to a handwritten Japanese essay by Student D achieved approximately 99% transcription accuracy using ChatGPT, as illustrated in Figure 4. While the digital text closely mirrors the handwritten version, a notable discrepancy is observed in the

transcription of the phrase “でした” (highlighted in red), where the OCR technology has incorrectly added unnecessary symbols (**). This inaccuracy likely resulted from the misinterpretation of handwriting characteristics, such as excessive spacing, overlapping strokes, or stray marks, which the OCR system misidentified as additional characters. Such minor but significant deviations highlight the need for careful oversight when using AI-driven OCR for educational purposes, ensuring that automated transcriptions are verified for accuracy before being used for feedback or assessment. This example illustrates that even with a high overall accuracy rate, minor recognition errors can occur, potentially altering the intended meaning of the student’s text, thus emphasising the importance of maintaining instructor validation in AI-assisted assessment workflows.

Essay of Student D

The student chose the topic 私のたんじょうび



わたしのたんじょうびはじゅうにに
じゅうよろこびです。わたしのとも
だちはたのしいと思い**、たのし
んでテレビをみました**。コーヒーを
のみました。のみから次の日**、と
もだちはおくりのろびやくろじゅう
ろくえんでした**。

~99% match

Figure 4. Essay by Student D and its transcribed version by ChatGPT

Source: ChatGPT

The findings indicate that ChatGPT performs well in identifying a range of surface-level language errors in handwritten Japanese essays, particularly those related to grammar, vocabulary use, and sentence structure. For instance, the tool successfully flags incorrect verb conjugations, such as *たべますでした* instead of the correct *たべました*, or the misuse of particles, such as *が* in place of *は* (e.g. *わたしが学生です* instead of *わたしは学生です*). It also detects the overuse of basic expressions such as *たのしい* or *おいしい* without lexical variation, encouraging students to diversify their vocabulary. Similarly, insufficient use of kanji (e.g. writing *たんじょうび* entirely in hiragana instead of the expected *誕生日*) is identified as an area for improvement. These capabilities demonstrate the model’s utility in providing consistent and timely feedback on recurring linguistic patterns, particularly in beginner- and intermediate-level writing.

However, the study reveals several limitations in ChatGPT’s ability to accurately assess context-dependent or stylistically flexible language use. In some cases, the model incorrectly flags idiomatic or casual expressions as errors, such as *じゃあね* or *よろしくね*, even when they are appropriately used in informal narratives. Moreover, OCR-related errors result in the misreading of handwritten characters; for example, *ぬ* may be misrecognised as *め*, or the kanji *話* might be transcribed incorrectly due to stroke overlap or spacing inconsistencies. Such misreadings can lead to false feedback or overlooked mistakes, highlighting

the risk of relying solely on automated transcription without human verification. These findings emphasise the continued need for educator supervision to validate AI output and maintain alignment with the learner’s intended meaning and writing style.

Beyond structural accuracy, ChatGPT also demonstrates a limited ability to assess higher-order writing features, such as tone, creativity, and cohesion. For example, it may not recognise inconsistencies in register within a single essay (e.g. switching between *です・ます* form and plain form: *うれしかったです*. *でも、ねむい*), or fail to comment on the appropriateness of emotional expressions in contextually rich writing. In narrative essays such as *私のたんじょうび*, students may include personal reflections, humour, or cultural references that the AI cannot meaningfully interpret or evaluate. Additionally, the tool does not assess coherence or flow between paragraphs, which are critical in longer written compositions. These limitations suggest that while ChatGPT enhances efficiency in detecting formal errors, it should function as a supplementary resource to human assessment, ensuring that language learning remains both accurate and pedagogically holistic.

Unexpectedly, while OCR-assisted transcription achieved high accuracy rates (approximately 90%-99%) in most cases, notable discrepancies were observed in instances involving irregular handwriting styles or poorly spaced characters, such as the misreading of *でした* or the misidentification of *ぬ* as *め*. These errors underscore

the limitations of OCR in handling nonstandardised scripts and suggest that visual noise – such as overlapping strokes or stray marks, can distort recognition accuracy. Moreover, ChatGPT occasionally flagged idiomatic or informal expressions (e.g. じゃあね, よろしくね) as incorrect despite their appropriateness in narrative contexts, revealing a lack of contextual and stylistic sensitivity. These findings may be attributed to the model's training on formal text corpora and insufficient exposure to learner-specific or culturally embedded language usage, thereby highlighting the need for human oversight in interpreting nuanced writing features.

This study demonstrates that ChatGPT, combined with OCR technology, can assist in identifying common language errors in handwritten Japanese essays, making it a useful tool for language instructors. It offers a practical method for quickly converting handwritten texts into digital form and providing students with individualised feedback. This approach is particularly beneficial in large or hybrid classes, where instructors may not have the capacity to offer detailed feedback to every student. At the same time, the results underscore the limitations of AI, such as difficulties with informal language or unclear handwriting, reinforcing the importance of teacher review to ensure accurate and meaningful feedback.

Based on the findings of this study, several policy-level recommendations can be made to support the effective and ethical integration of AI tools in language education. First, institutions should consider incorporating AI-assisted assessment technologies – such as ChatGPT in combination with OCR – into their official teaching and evaluation frameworks. This integration can enhance assessment efficiency and improve the quality of feedback. To ensure consistency and accountability, it is essential to develop institutional guidelines that define the scope of AI use, including when and how AI-generated feedback should be reviewed or supplemented by instructors. Second, data privacy and student consent protocols must be established and aligned with national and institutional data protection standards. Given that student work is uploaded and processed through AI systems, policies should clearly address the storage, sharing, and deletion of student-generated content to safeguard personal information.

Finally, professional development programmes should be offered to educators to strengthen their digital literacy and equip them with the skills needed to critically evaluate and interpret AI-generated outputs. This will help maintain the balance between automation and pedagogical intent, ensuring that AI remains a support tool rather than a replacement for human feedback in educational assessment. While the study demonstrates the potential of ChatGPT and OCR technology in supporting language assessment, several gaps remain. Notably, the system struggles with informal expressions, creative writing elements, and varied handwriting styles, raising questions about its applicability across diverse learner populations and writing genres. This study demonstrates that ChatGPT holds significant promise

as a support tool for language educators assessing handwritten Japanese essays, particularly at beginner and intermediate levels. As no existing studies were found that directly explore ChatGPT's application to handwritten Japanese as a foreign language, this research addresses a clear gap in AI-assisted language assessment and contributes original evidence from a linguistically underrepresented context.

The institutional push towards AI-enhanced formative assessment is further evidenced in national policy developments. Singapore's National AI Strategy 2.0 prioritises the development of human-centric AI applications in education to improve instructional effectiveness and learner outcomes (Smart Nation and Digital Government Office, 2023). In alignment with this, the Ministry of Education (2023) advocates for the responsible use of AI to support differentiated instruction and reduce teacher workload.

In the evolving field of Japanese language education, researchers have been exploring the integration of AI and OCR technologies to improve the assessment and support of student writing. Y. Jing (2020) highlighted the potential of natural language processing (NLP) in personalising learning experiences and enhancing teaching efficiency in Japanese instruction. Similarly, W. Huang (2020) developed a machine learning-based scoring system focused on lexical and grammatical features to assist learners in developing their writing proficiency. P. Taele *et al.* (2025) introduced the “Kanji Workbook”, an intelligent tutoring system designed to simulate teacher feedback and guide students in mastering kanji characters. These innovations reflect a growing recognition of AI's role in supplementing traditional language instruction through automated support tools. Building on this trajectory, the present study extends the discussion by applying ChatGPT to handwritten Japanese essays – an area previously underexplored. Unlike systems that rely solely on rule-based scoring or fixed feature analysis, ChatGPT provides dynamic, context-sensitive feedback, supporting both grammatical accuracy and learner reflection. By evaluating how ChatGPT responds to authentic handwritten texts, this research offers pedagogically meaningful insights and underscores the importance of balancing automation with instructional sensitivity in AI-mediated language learning environments.

Building on the conceptual research of J. Li *et al.* (2024), which outlined ChatGPT's potential in second language (L2) writing – from feedback and assessment to collaborative learning – this study offers classroom-based validation. It affirms J. Li's vision of ChatGPT as a feedback and assessment tool while advancing the discourse by examining how the model performs with handwritten input. Notably, it identifies limitations in ChatGPT's capacity to interpret stylistic nuance and context, thereby contributing a balanced and practice-oriented perspective to the growing field of AI-enhanced language education.

Other recent studies support the relevance of ChatGPT across broader educational domains. T. Aydin Yıldız (2024) demonstrated its role in fostering 21st-century teaching skills, while the present study shows its utility from

the learner's perspective, reinforcing ChatGPT's value across the teaching-learning spectrum. Similarly, J. Ilić *et al.*'s (2024) review of AI in higher STEM education highlights personalisation and engagement – two features echoed in the feedback capabilities observed in this research. Technologically, this study differs from the JESS system developed by T. Ishioka & M. Kameda (2004), which scores Japanese essays based on fixed metrics such as structure, vocabulary complexity, and passive constructions, independent of human input. In contrast, ChatGPT offers dynamic, interactive feedback, functioning not as a replacement for educators but as a supplementary tool to support formative assessment and reflective learning.

C. Tsai's (2024) study using convolutional neural networks (CNNs) to recognise Japanese characters complements this study by addressing another facet of handwritten text processing. While Tsai emphasises technical recognition accuracy, the present research approaches handwritten language from an instructional feedback perspective. Together, these studies highlight the multifaceted benefits of AI: precise character-level recognition and meaningful, pedagogically relevant feedback at the learner level. Thus, this study aligns with concerns raised by O. ZawackiRichter *et al.* (2024), who critiqued the dominance of STEM-driven, theory-light AIED research. By presenting a theory-informed, classroom-tested use case of ChatGPT, this research exemplifies how AI can be thoughtfully integrated into pedagogical practice. It underscores the importance of maintaining a learner-centred and educator-guided approach to AI deployment – one that enhances rather than replaces human interaction.

Therefore, this study builds on prior research by demonstrating ChatGPT's potential as a valuable tool for educators in assessing handwritten Japanese essays, providing dynamic, contextsensitive feedback that surpasses the limitations of rule-based systems. It highlights how ChatGPT can support teachers in delivering timely, formative insights while preserving the educator's central role in guiding student learning and interpretation.

CONCLUSIONS

The study found that ChatGPT is a valuable tool for analysing students' handwritten essays in Japanese, offering both individualised and cohort-level feedback that can support more targeted and efficient language instruction. Through its evaluation, ChatGPT consistently identified eight recurring categories of language difficulties: limited or incorrect use of kanji, grammatical errors (especially with particles and verb tenses), basic and repetitive sentence structures, narrow vocabulary range, spelling and phonetic transcription errors, redundancy, incoherent phrasing, and inappropriate contextual usage. Quantitative analysis showed that the most frequent issues were the lack of kanji (25%), overuse of basic sentence structures (20%), and limited vocabulary (15%). These

were followed by repetition (10%), grammatical issues (10%), misused words (10%), tense inconsistency (5%), and redundancy (5%). ChatGPT's ability to highlight these patterns enabled teachers not only to address individual learning gaps but also to make data-informed decisions on adjusting instruction at the group level. Additionally, ChatGPT's rapid generation of error summaries and suggested improvements helped reduce the time educators spent on marking, while enhancing the quality and consistency of feedback. Teachers were able to track common trends, observe progression across drafts, and design supplementary exercises targeting frequently occurring issues. However, the research also noted limitations: ChatGPT was not capable of assessing students' emotional engagement, creativity, or psychological factors, all of which can influence language performance. Therefore, the integration of AI proved most effective when combined with teachers' professional judgement and pedagogical expertise.

Overall, the findings suggest that ChatGPT can significantly optimise the essay feedback process, enhance learners' awareness of their own mistakes, and contribute to more personalised and responsive language education when used as a complementary tool rather than a replacement for human instruction. The results underscore the pedagogical potential of generative AI to support formative assessment through personalised, timely feedback, while also prompting critical reflection on the role of human-AI collaboration in instructional design. Notably, the study identifies key technological limitations, such as OCR sensitivity to handwriting variation, which may affect transcription accuracy and, consequently, the reliability of assessment. Ethical considerations surrounding student data privacy and the potential for pedagogical over-reliance on AI further emphasise the necessity of instructor oversight. In addition, the research highlights the imperative of strengthening digital competence among educators to ensure the informed and responsible adoption of AI tools. Despite promising results, further research is required to better understand ChatGPT's effectiveness and limitations in supporting educators, including its performance across various languages and scripts, the long-term impact of AI-assisted feedback, and user perceptions compared to human evaluation. Interdisciplinary efforts are also needed to improve OCR accuracy, integrate AI into institutional assessment frameworks, and ensure alignment with ethical, legal, and pedagogical standards.

ACKNOWLEDGEMENTS

None.

FUNDING

None.

CONFLICT OF INTEREST

None.

REFERENCES

- [1] Aydın Yıldız, T. (2024). Exploring the impact of ChatGPT on improving 21st-century skills for future English teachers during lesson planning. *Computers in the Schools*, 1-24. doi: [10.1080/07380569.2024.2429534](https://doi.org/10.1080/07380569.2024.2429534).
- [2] Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62. doi: [10.61969/jai.1337500](https://doi.org/10.61969/jai.1337500).
- [3] Chaudhry, I.S., Sarwary, S.A.M., El Refae, G.A., & Chabchoub, H. (2023). Time to revisit existing student's performance evaluation approach in higher education sector in a new era of ChatGPT – a case study. *Cogent Education*, 10(1), article number 2210461. doi: [10.1080/2331186X.2023.2210461](https://doi.org/10.1080/2331186X.2023.2210461).
- [4] Dwivedi, Y.K., et al. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, article number 101994. doi: [10.1016/j.ijinfomgt.2019.08.002](https://doi.org/10.1016/j.ijinfomgt.2019.08.002).
- [5] Huang, W. (2020). Machine learning approach of Japanese composition scoring and writing aided system's design. *arXiv Preprint*, 225-230. doi: [10.48550/arXiv.2008.11488](https://doi.org/10.48550/arXiv.2008.11488).
- [6] Ilić, J., Ivanović, M., & Klačnja-Milićević, A. (2024). The impact of ChatGPT on student learning experience in higher STEM education: A systematic literature review. In *2024 21st International conference on information technology based higher education and training (ITHET)* (pp. 1-9). Paris: IEEE. doi: [10.1109/ITHET61869.2024.10837649](https://doi.org/10.1109/ITHET61869.2024.10837649).
- [7] Ishioka, T., & Kameda, M. (2004). Automated Japanese essay scoring system: Jess. In *Proceedings of the 21st international conference on computational linguistics and 44th annual meeting of the association for computational linguistics* (pp. 4-8). Zaragoza: IEEE. doi: [10.1109/DEXA.2004.1333440](https://doi.org/10.1109/DEXA.2004.1333440).
- [8] Jing, Y. (2020). Research on the application of artificial intelligence natural language processing technology in Japanese teaching. *Journal of Physics: Conference Series*, 1682, article number 012081. doi: [10.1088/1742-6596/1682/1/012081](https://doi.org/10.1088/1742-6596/1682/1/012081).
- [9] Kamoun, F., El Ayeb, W., Jabri, I., Sifi, S., & Iqbal, F. (2024). Exploring students' and faculty's knowledge, attitudes, and perceptions towards ChatGPT: A cross-sectional empirical study. *Journal of Information Technology Education: Research*, 23, article number 4. doi: [10.28945/5239](https://doi.org/10.28945/5239).
- [10] Kotsis, K.T. (2024). Correcting students' misconceptions in physics using experiments designed by ChatGPT. *European Journal of Contemporary Education and E-Learning*, 2(2), 83-100. doi: [10.59324/ejceel.2024.2\(2\).07](https://doi.org/10.59324/ejceel.2024.2(2).07).
- [11] Li, J., Jangamreddy, N.K., Hisamoto, R., Bhansali, R., Dyda, A., Zaphir, L., & Glencross, M. (2024). AI-assisted marking: Functionality and limitations of ChatGPT in written assessment evaluation. *Australasian Journal of Educational Technology*, 40(5), 56-72. doi: [10.14742/ajet.9463](https://doi.org/10.14742/ajet.9463).
- [12] Li, M. (2024). Leveraging ChatGPT for second language writing feedback and assessment. *International Journal of Computer-Assisted Language Learning and Teaching*, 14(2), 1-11. doi: [10.4018/IJCALLT.360382](https://doi.org/10.4018/IJCALLT.360382).
- [13] Ministry of Education. (2023). Retrieved from <https://www.moe.gov.sg>.
- [14] Mohan, G.B., Kumar, R.P., Krishh, P.V., Keerthinathan, A., Lavanya, G., Meghana, M.K.U., Sulthana, S., & Doss, S. (2024). An analysis of large language models: Their impact and potential applications. *Knowledge and Information Systems*, 66, 5047-5070. doi: [10.1007/s10115-024-02120-8](https://doi.org/10.1007/s10115-024-02120-8).
- [15] Pesce, M.A., & Fernández Blanco, D. (2024). Perceptions of pre-service teachers on the use of ChatGPT during their training and the future implications for their future role in high schools: A research in progress. *Research Square*, 1-28. doi: [10.21203/rs.3.rs-4542918/v1](https://doi.org/10.21203/rs.3.rs-4542918/v1).
- [16] Pham, M.T., & Cao, T.X.T. (2025). The practice of ChatGPT in English teaching and learning in Vietnam: A systematic review. *International Journal of TESOL & Education*, 5(1), 50-70. doi: [10.54855/ijte.25513](https://doi.org/10.54855/ijte.25513).
- [17] Smart Nation and Digital Government Office. (2023). Retrieved from <https://file.go.gov.sg/nais2023.pdf>.
- [18] Susnjak, T., & McIntosh, T.R. (2024). ChatGPT: The end of online exam integrity?. *Education Sciences*, 14(6), article number 656. doi: [10.3390/educsci14060656](https://doi.org/10.3390/educsci14060656).
- [19] Taele, P., Koh, J.I., & Hammond, T. (2020). Kanji workbook: A writing-based intelligent tutoring system for learning proper Japanese kanji writing technique with instructor-emulated assessment. In *Proceedings of the AAAI conference on Artificial Intelligence* (vol. 34(08), pp. 13382-13389). Washington: AAAI Press. doi: [10.1609/aaai.v34i08.7053](https://doi.org/10.1609/aaai.v34i08.7053).
- [20] Tsai, C. (2016). *Recognizing handwritten Japanese characters using deep convolutional neural networks*. Retrieved from https://cs231n.stanford.edu/reports/2016/pdfs/262_Report.pdf.
- [21] Woodruff, K., Hutson, J., & Arnone, K. (2023). Perceptions and barriers to adopting artificial intelligence in K-12 education: A survey of educators in fifty states. In *Reimagining education: The role of e-learning, creativity, and technology in the post-pandemic era*. InTechOpen. doi: [10.5772/intechopen.1002741](https://doi.org/10.5772/intechopen.1002741).
- [22] Zawacki-Richter, O., Marín, V.I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators?. *International Journal of Educational Technology in Higher Education*, 16(1), article number 39. doi: [10.1186/s41239-019-0171-0](https://doi.org/10.1186/s41239-019-0171-0).

Марія Щедріна

Аспірант, викладач

Інститут вищої освіти Національної академії педагогічних наук України

01014, вул. Бастіонна, 9, м. Київ, Україна

Наньянська політехніка

569830, просп. Анг Мо Кіо, 180, м. Сінгапур, Сінгапур

<https://orcid.org/0000-0002-0214-2024>

Юе Ман Джуді Іп

Старший викладач

Наньянська політехніка

569830, просп. Анг Мо Кіо, 180, м. Сінгапур, Сінгапур

<https://orcid.org/0009-0002-3377-6609>

Оцінювання ChatGPT у викладанні японської мови та перевірці есе

Анотація. Актуальність дослідження зумовлена зростаючою потребою у впровадженні штучного інтелекту в освітній процес, зокрема у викладанні мов із нелатинським письмом, як-от японська. Поєднання оптичного розпізнавання символів і аналізу на базі ШІ відкриває нові можливості для підвищення ефективності перевірки письмових робіт, якості зворотного зв'язку та індивідуалізації навчання. Метою дослідження було вивчення можливостей використання ChatGPT для підвищення ефективності викладання японської мови шляхом автоматизованої перевірки студентських есе. У методологічному плані викладачі використали ChatGPT для оцифрування рукописних есе студентів, що дозволило проаналізувати граматику, лексику, структуру речень і вживання канджі. Було застосовано змішаний підхід: кількісний аналіз типових помилок поєднано з якісною оцінкою викладачів. Результати дослідження показали, що ChatGPT ефективно ідентифікує вісім основних категорій помилок, зокрема недостатнє використання канджі, надмірне застосування базових граматичних конструкцій і обмежений лексичний запас. Здатність інструмента створювати миттєві резюме помилок і надавати цільовий зворотний зв'язок суттєво зменшила час перевірки, забезпечуючи послідовну та персоналізовану підтримку студентів. Результати свідчать про те, що ChatGPT може стати цінним доповненням до традиційних методів викладання, покращуючи якість зворотного зв'язку, підтримуючи диференційоване навчання та сприяючи автономії здобувачів освіти. Отримані висновки можуть стати корисними для викладачів мов, розробників навчальних програм і фахівців з освітніх технологій, які прагнуть інтегрувати штучний інтелект у викладання та оцінювання рідковживаних мов, зокрема японської

Ключові слова: генеративний штучний інтелект; оцінювання мовних знань; ШІ в освіті; автоматизований зворотний зв'язок; методика викладання японської мови