ChatGPTと英語教育の統合に向けて
──日本の理系大学における予備調査と授業設計からの示唆──
Integrating ChatGPT Into English Education: Insights From a Preliminary Survey and Course Design at a Japanese Science University

宮添　輝美 MIYAZOE, Terumi
● 東京理科大学
Tokyo University of Science

ABSTRACT

　2022年11月に登場した生成AIは，従来の教育方法を根底から覆す「破壊的テクノロジー」となるだろう。本稿では，生成AI言語モデルを利用した英語授業設計を目指し，学生に対して実施した予備調査とその結果に基づく新しい授業設計案を報告する。学生の英語学習におけるAI校閲ツールとChatGPTの使用状況，ならびにこれらに対する認識を調査し，得られたデータを基に2023年春学期の授業設計を修正した。2023年4月時点の調査では，多くの学生が既にAIを英文校正のツールとして利用しており，ChatGPTの多様な活用可能性を認識していた。教育界は，この新しい技術の可能性を認め，言語教育の旧来の方法を見直す必要がある。本研究は，生成AIの登場により変わりつつある現代の高等教育の現場を捉え，実践と研究の連動により迅速に授業設計するプロセスを例示し，高等教育におけるAIツールの利活用法を提案する。

　Introduced in November 2022, generative AI has emerged as a potential 'disruptive technology,' posing a necessity for a comprehensive paradigm shift in education globally. This paper presents a preliminary survey that probes into the perceptions and applications of ChatGPT, a generative AI language model, by students aiming to enhance their English proficiency at a Japanese science university. Additionally, it explores the recent advancements of ChatGPT in English education, yielding practical insights for its seamless integration into language instruction. The study examines students' engagement and attitudes towards AI editing tools and ChatGPT in their pursuit of English studies. The survey, conducted in April 2023, captures the students' initial responses to this nascent technology, revealing a substantial number of them already incorporating AI editing in their writing practices and acknowledging its extensive potential applications. This paper elucidates the decision-making process integral to language course design, highlighting the revolutionary impact of advanced AI on education. It stands as a testament to the synergistic relationship

between research and pedagogical practices, offering strategic guidance for the proficient integration of AI-driven language tools in higher education.

1. Introduction

The education landscape in the post-COVID era is facing a new challenge: how to adapt to the unprecedentedly sophisticated artificial intelligence exemplified by ChatGPT. Language education is particularly affected (Felten et al., 2023), given ChatGPT's exceptional capability to handle multilingual data across various disciplines, incomparable to the traditional AI editing tools focused on basic syntax, spelling, and grammar checking. Yet, as of the start of the 2023 academic year in April, there had been little discussion regarding the implications of generative AI in this field. This study aims to address that gap, seeking to understand student perceptions and potential educational strategies. It captures the pre-implementation sentiments—a snapshot in time that would be impossible to replicate—set against the backdrop of the ongoing Fourth Industrial Revolution (World Economic Forum, 2023) that is influencing every facet of our lives and learning.

1.1 An Introduction to Generative AI (ChatGPT)

ChatGPT, developed by OpenAI, stands as a prime example of sophisticated conversational language models that generate human-like text. Since its public release on November 30, 2022 (OpenAI, 2022), ChatGPT has set several benchmarks, indicating its swift adoption in various sectors and everyday life (Business of Apps, 2023). The evolution from the original GPT to the most recent version, GPT-4, is detailed in three notable technical reports (Brown et al., 2020; OpenAI, 2023a; Radford et al., 2019).

GPT, short for Generative Pre-trained Transformer, was the initial model primed to generate human-like text across a plethora of domains and disciplines. This encompassed tasks like reading, translation, summarization, and answering questions. Its ability to multitask was honed by training the model on a vast corpus of WebText data, enabling GPT to manage new tasks even if not directly trained on them—a feature known as 'zero-shot learning.'

GPT-2, the subsequent model, was a refinement of the original, though it still occasionally churned out nonsensical sentences. GPT-3 rectified these issues and flaunted enhanced multitasking capabilities, producing more contextually accurate outputs. By utilizing a 'few-shot learning' approach, GPT-3 could adapt to diverse tasks. This method, 'in-context learning,' was underpinned by an expansive training dataset and required minimal fine-tuning. Additionally, GPT-3 employed 'autoregressive' probability calculations for predicting subsequent words in a sequence, resulting in highly authentic human-like text.

GPT-4, the latest in the lineup, amplifies the model's accuracy and introduces multimodal capabilities, allowing it to process both visual and textual data. Its proficiency mirrors human-like standards in various professional and academic spheres. OpenAI dedicated significant sections of the GPT-4 report to the "System Card," outlining the measures to curb potential misuse and risks. As for language education, Figure 5 (OpenAI, 2023a, p. 8) of the GPT-4 report elucidates its multilingual prowess using the MMLU or Multitask Language Understanding benchmark, highlighting its dominant performance in English relative to other languages.

The ChatGPT series has undergone rapid refinements, evolving impressively into an unprecedented tool for various purposes. As advancements continue to emerge, the field of

educational research is progressively adapting, making strides to integrate this innovative technology into teaching methodologies. This integration, particularly in the domain of language education, will be further explored in the subsequent sections.

1.2 A Review of Generative AI Applications in English Education

As of the end of July 2023, six months following the launch of ChatGPT, a modest number of studies focusing on its application in English education have been identified, as indexed in the Scopus database—one of the most extensive abstract, citation, and indexing databases, widely used by researchers, academic institutions, and government agencies (Elsevier, n.d.). Reflecting the early stage of this generative AI technology, the existing research is largely exploratory and varies in its approach to utilizing ChatGPT. Despite the limited timeframe and the inherent constraints in academic publishing, these studies can be broadly categorized into three thematic areas: ChatGPT's role in vocabulary and reading development, its use for specific language purposes, and perspectives from the faculty.

In the realm of vocabulary and reading development, Ehara (2023) in Japan presents an exploratory system through the GPT's natural text-generation ability to create personalized adaptive learning experiences in English education. In his proposed system, learners initially rate their familiarity with each of the 12,000 English words on a five-point Likert scale and indicate their preferred topic domain. Based on this data, GPT generates reading texts, each incorporating three words that are predicted to be unfamiliar to the learner, thereby facilitating targeted vocabulary acquisition. This usage could be applied to a wide range of uses within language teaching and learning, each tailored to the precise needs of the individual learner. Kohnke et al. (2023), in their Hong Kong-based study, relied on the AI's multilingual ability (in this case, Chinese-English translation) to serve as a tailored dictionary offering example sentences for vocabulary enhancement and note-making, its adeptness at producing natural text for various purposes with different styles and tones, its adaptability in adjusting language complexity to meet specific goals, and its ability to generate comprehension questions from reading texts, complete with explanations. The study also explores potential drawbacks and ethical issues surrounding this burgeoning technology, including occasional inaccuracies, possible cultural biases, and its impact on the assessment process.

In the context of specific language purposes, Kovačević (2023) from Bosnia and Herzegovina identifies four key areas where ChatGPT can be applied, which are applicable across a wide range of ESP specializations: (a) generating texts relevant to specific fields of study; (b) creating vocabulary and grammar exercises; (c) serving as virtual tutors or chatbots; and (d) providing real-time feedback and evaluations. While acknowledging the pre-existing capabilities of ChatGPT, Kovačević posits that the creation of a domain-specific dataset for fine-tuning could enhance the model's performance and effectiveness in these educational contexts. Similarly, Young and Shishido (2023) explore the potential use of ChatGPT-generated English dialogues as a resource for English language practice materials. By employing three distinct readability scales—the Flesch reading ease, Dale-Chall readability formula, and MacAlpine EFLAW readability formula—they assess the complexity and understandability of the dialogues generated to conclude that the dialogues fall within the CEFR A2 (elementary) and B1 (intermediate) levels of English proficiency (Council of Europe, n.d.), thereby making ChatGPT a suitable tool for creating a wide range of learning materials.

From the perspective of faculty, Mohamed (2023) at Northern Border University in Saudi Arabia is different from the previous studies in that it conducts in-depth interviews with ten EFL faculty members to explore the benefits, challenges, and potential applications of ChatGPT in EFL instruction. His comprehensive literature review efficiently summarizes the advantages (personalized learning, immediate feedback, accessibility) and drawbacks (lack of human interaction, potential bias and discrimination, limited language exposure) of AI-assisted language learning (prior to the advent of the generative AI) in previous research. The interview results provide a wealth of insights into EFL faculty members' diverse viewpoints on this emerging technology, highlighting important considerations such as ethical and privacy concerns.

1.3 Research Questions

The literature demonstrates that as of Spring 2023, while the potential of this novel technology in teaching and learning was being explored, its full impact was still unfolding. Notably, Young and Shishido's (2023) observation of English dialogue levels was based on their prompts, overlooking ChatGPT's ability to fine-tune to specific English proficiency levels. The current study was undertaken to integrate a new tool—generative AI—into existing course designs that prominently feature essay or script writing. This integration serves as a baseline for understanding the impact of generative AI in educational contexts. In doing so, this study aims to contribute to the ongoing academic conversation by formulating two research questions that guide the exploration of innovative course designs utilizing this new technology:

1. To what extent are students familiar with traditional AI editing tools and emergent generative AIs, such as ChatGPT, in their English studies? (RQ 1)

2. Which strategies prove most effective in integrating AI tools into English courses across different proficiency levels? (RQ 2)

The first question was established to gauge the feasibility of using generative AI in course designs, while the second aims to pinpoint specific instructional strategies suitable for diverse educational settings.

1.4 Research Framework of the Current Study

The trajectory of this study is informed by the ADDIE model (Gagne et al., 2005), which underscores the importance of iterative improvement in instructional design. In this vein, the survey acts as a needs analysis, assessing students' readiness for the inclusion of AI tools in their learning journey. The research insights will inform the integration of these tools in the upcoming course designs, setting the stage for subsequent rounds of evaluations and refinements in instructional design.

Specifically, Section 4, 'Results,' directly addresses and provides answers to RQ1, elucidating the insights obtained from the survey conducted for this study. Conversely, Section 5, 'Impact on Course Design,' is designed to offer responses to

Figure 1
Interplay of Research and Practice Based on ADDIE Model
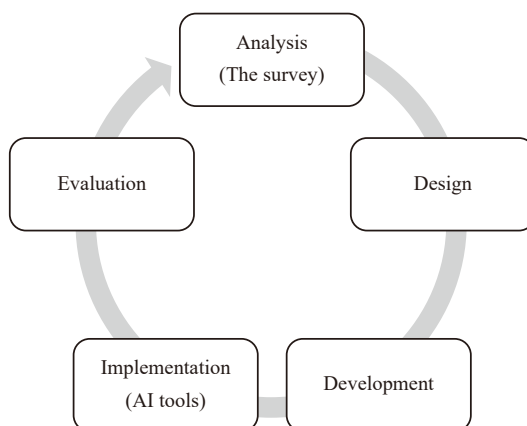
RQ2, showcasing a revised course design that judiciously incorporates AI tools.

2. Method

A survey was conducted at the outset of the spring semester in 2023. It targeted students enrolled in five distinct English language classes, encompassing both students already registered for the course and those deliberating during the course guidance week.

The survey aimed to gauge students' engagement with and perceptions of AI tools, quantitatively and qualitatively (Creswell, 2003). Incorporating specific AI-related questions aimed not only to glean insights but also to inform students about potential AI integration in forthcoming courses. This approach ensured that students apprehensive about AI could explore alternative courses led by different instructors.

Participation in the survey was voluntary. The survey was facilitated via the online platform, Quartics, an officially licensed survey-specific research tool of the university. Accessible from April 11 to 21, students received bi-weekly oral reminders. The researcher diligently documented all related activities to uphold the integrity of the research.

The survey was composed to explore students' perceptions and utilization of AI English editing tools, notably ChatGPT, Microsoft Word editing, Grammarly, Writefull, and ProWritingAid. Given that the university licenses Microsoft Office 365, all students can access Word. Nevertheless, the frequency of using its English editing features remained ambiguous. The questionnaire encompassed two multiple-choice questions on English learning attitudes, four on AI tool utilization and perceptions, four demographic inquiries, and two open-ended questions. While the primary survey was in Japanese for accessibility, an English translation is attached for reference.

For analysis, quantitative data underwent using SPSS for intricate statistical evaluations and EXCEL for visualization. Qualitative data were assessed using a coding method anchored in grounded theory (Corbin & Strauss, 1990). A text-mining AI tool was employed for triangulation, ensuring the robustness of results.

The university where this research was conducted mandates specialized ethics reviews specifically for research that could cause potential medical or biological harm to humans. Even so, this survey was carried out with voluntary participation. Participants were clearly informed on the survey's initial page about its purpose, the measures taken to ensure data privacy, and the potential use of their anonymized data for presentations and publications.

It's important to note that due to beginning-of-term dynamics and the imperative to swiftly adjust course plans based on insights, preliminary data was immediately accessible. The real-time feedback, rendered as visual charts in the Quartics system, facilitated rapid decisions regarding course design modifications.

3. Results

A total of 130 responses were received. After removing 10 responses, 120 remained for quantitative analysis. These removals accounted for incomplete submissions, including one case of double entry which was possibly due to technical issues. The two open-text questions, Question 3.2 and Question 5 (hereafter, Q3.2 and Q5), yielded 114 and 77 comments respectively. These comments amounted to approximately 3,460 and 2,290 characters each in Japanese. Minimum modifications were made to the comments to ensure consistency, especially given the Japanese language's use of different scripts to represent the

same words.

3.1 Respondents' Profile

The survey participants (N=120) were undergraduate students from a science-focused university located in central Tokyo. They were enrolled in English courses of various CEFR levels (Council of Europe, n.d.), from A2 to C1. These courses were to be taught by the author. The distribution of the respondents across the courses was as follows: Advanced (27 students), Intermediate (83 students), and Basic (17 students). Among these, seven students contemplated enrolling in two courses by the author. However, participation in this survey was limited to once per student.

The academic disciplines of the respondents spanned five fields: Mathematics (43 students, 35.8%), Physics (13, 10.8%), Chemistry (58, 48.3%), Architecture (2, 1.7%), and Management (4, 3.3%). The higher representation of Mathematics and Chemistry students might be due to the current courses being more readily available as options within their divisions compared to other disciplines. The gender distribution was as follows: males (87, 72.5%), females (28, 23.3%), and those who chose not to specify (5, 4.2%). This distribution closely mirrored the university's overall gender ratio of 77.4% male to 22.6% female (TUS, 2023).

Respondent ages ranged from 18 to 24 years. The majority were 18 or 19 years old (69 students, 57.5%), followed by those aged 20 or 21 (35 students, 29.2%). The remaining 13.3% (16 students) were older than 21, including working adult students.

3.2 Students' Use of AI Editing Tools

Q2 focused on AI editing tools. Sub-question Q2.1 sought information about the frequency of use, while Q2.2 asked about specific tools students had utilized.

Figure 2 reveals that roughly 40% of students had experience using AI editing tools. Within the context of this research, where students were introduced to these tools, there's a plausible assumption that they might continue using them.

As seen in Figure 3, among the 40% who regularly used AI tools, the majority relied on Word's editing functions, followed by Grammarly. It should be highlighted that Word's editing capabilities vary depending on its version. The latest Microsoft Editor, for instance, offers a host of advanced features, including similarity checks against web content. However, these functions might not be as specialized as those found in dedicated AI tools. In the open response section labelled 'Other (fill in)', students indicated a combinatory use of multiple AI editing tools: eight

Figure 2

Frequency of Students' Use of AI Editing Tools Before the Course (Q2.1)

	Frequency	Percent	Cumulative Percent
I use it frequently.	9	7.5	7.5
I use it occasionally.	29	24.2	31.7
I've used it once or twice.	13	10.8	42.5
I've heard of it but have never used it.	37	30.8	73.3
I've never heard of it (or don't know it).	32	26.7	100.0
Total	120	100.0	

Figure 3
Types of AI Editing Tools Students Used Before the Course (Q2.2)

Word editing tool		34	28.3	28.3
	Word+Grammarly	8	6.7	35.0
	Word+Google translate	2	1.7	36.7
Writefull		1	0.8	37.5
Grammarly		19	15.8	53.3
	Grammarly+ProWritingAid	1	0.8	54.2
	Grammary+DeepL	1	0.8	55.0
ProWritingAid		1	0.8	55.8
Google Translate		6	5.0	60.8
DeepL		4	3.3	64.2
ChatGPT		1	0.8	65.0
Not given		42	35.0	100.0
	Total	120	100.0	

students mentioned Google Translate, five DeepL—a free AI-based translation service—and one referred to ChatGPT. Some students might have misunderstood the question, confusing editing tools with translation tools, or they might be using translation tools in a unique way to aid their editing process.

3.3 Students' Use and Perception of ChatGPT

The Q3 section on ChatGPT use is structured as follows: Q3.1 inquires about the students' frequency of using ChatGPT; To avoid leading responses, Q3.2 was purposefully positioned before Q3.3. This arrangement provides an open-ended space for students to articulate their understanding or perception of ChatGPT without bias; and Q3.3 focuses on the purposes for which students might use ChatGPT.

Results from Q3.1 and Q3.3, due to their quantitative nature, are presented first, followed by the qualitative findings from Q3.2. Figure 4 captures the direct responses of all 120 students. Conversely, Figures 5 and 6, allowing for multiple

selections, result in totals surpassing the base sample size of 120. This design offers deeper insights into the students' diverse perceptions and uses of ChatGPT.

3.3.1 Students' Use of ChatGPT

Figures 4 and 5 depict the extent and purpose of students' use of ChatGPT at the start of the course in April 2023. While only 20% had experience with ChatGPT, those who did were already leveraging the tool for advanced consultation on content and structure, not just for basic grammar and spelling checks. Interestingly, the use spans across all four linguistic skills: listening, speaking, reading, and writing. This data suggests that certain students were not merely using ChatGPT as a writing aid, but as a comprehensive language tool.

The open-ended comments from Q3.2 revealed that ten students utilized ChatGPT in unique ways, spanning from leisurely activities like gaming and conversational engagement, to functional tasks such as schedule management, slang term checks, and transcription services. Some even tapped into

ChatGPT for intricate tasks like programming, math calculations, and proofs.

3.3.2 Students' Perceptions of ChatGPT

Figure 6 provides an organized view of the text comments from 114 students regarding their perceptions of ChatGPT. The analysis identified common themes across these comments. Even though these themes were often expressed using varied terminology, they typically conveyed similar meanings. The emergence of these themes in the comments is quantified as 'Frequency 1.'

These comments were further classified based on their nature and focus into three primary categories: (a) responses and assessments related to ChatGPT; (b) descriptions of how they utilized the tool; and (c) other miscellaneous comments. These are delineated in 'Frequency 3.'

To provide a more nuanced analysis, we further subdivided these primary categories into sub-categories. For instance, comments under the 'Responses and Assessments' category were classified based on their sentiment into negative ('a') and positive ('b') evaluations. Other sub-

Figure 4

Frequency of Students' Use of ChatGPT Before the Course (Q3.1)

	Frequency	Percent	Cumulative Percent
I use it frequently.	4	3.3	3.3
I use it occasionally.	8	6.7	10
I've used it once or twice.	15	12.5	22.5
I've heard of it but have never used it.	70	58.3	80.8
I've never heard of it (or don't know it).	23	19.2	100
Total	120	100	

Figure 5

Purpose of Students' Use of ChatGPT Before the Course (Q3.3)

Purpose of use	Frequency	Percent	Cumulative Percent
English proofreading (grammar/spell check)	14	10.3	10.3
Content consultation and advice	14	10.3	20.6
Composition and structuring of sentences	9	6.6	27.2
Writing style check	6	4.4	31.6
Reading support	3	2.2	33.8
Listening support	3	2.2	36
Speaking assistance	2	1.5	37.5
Other (fill in)	10	7.4	44.9
Not used	75	55.1	100
Total	136	100	

Note. Students were allowed to select multiple options.

Figure 6

Students' Perceptions of ChatGPT Before the Course (Q3.2)

Categories	Themes	Frequency 1	Frequency 2		Frequency 3
Responses and Assessments	Doubtful/Mistakes	16	a	19	
	Underdevelopment	3			
	Accurate	4			32
	Convenient/Useful	6	b	13	
	Amazing/Intelligent	3			
Usage	Conversation/Dialogue Partner	7	c	10	
	Advice	3			
	Programming	4	d	4	
	Correction/Proofreading	4			
	Summarizing	2		11	70
	Translation	5	e		
	Writing Papers/Reports	7		18	
	Creation/Creativity	11			
	Answering Questions	20	f	27	
	Search/Information Provision	7			
Others	Not Sure	20		20	
	Others	24	g	24	
Total					146

Note. Responses could contain multiple themes.

categories encompassed references to ChatGPT's presumed emotional response ('c'), its computational proficiencies ('d'), its adeptness in offering writing support ('e'), its capacity for static information exchange ('f'), and other assorted observations ('g'). The counts for these subcategories are represented in 'Frequency 2.'

The student responses showcased a dichotomy of feelings and evaluations toward generative AI tools. Remarkably, prior to the course, many had begun experimenting with ChatGPT for diverse purposes, ranging from information retrieval to specialized aids in programming and writing. Some students viewed ChatGPT as a confidant, turning to it for advice on personal and sensitive matters.

Responses and assessments related to ChatGPT:

S42: "I know that ChatGPT is a type of AI development using deep learning and other methods provided by OpenAI. This AI can be used in a wide variety of ways, from creative tasks such as writing novels and thinking up names for people and pets, to systematic ones such as asking about the usage of software and apps, or assisting in programming development. As for its use in English classes, I think it's technically possible to use it for translation, of course, and even to write reports based on specified themes and word counts."

S75: "It seems amazing and fun. I don't think it's

educationally good to ask it to do all the assignments and copy them verbatim, but in the future information society, search ability will be necessary, so the ability to use it is also necessary. Until now, when researching specialized subjects, it was difficult for beginners to read because Wikipedia and the like use a lot of technical terms. But I thought it was a good feature that when you ask ChatGPT to 'explain a technical matter in a way that is easy for an elementary school student to understand,' it explains it in a simple way. Academia is becoming more diversified and specialized, so I think it's okay to ask ChatGPT about the touchpoints of a field of study you're just starting to learn about."

S26: "I recognize the need to develop the skill of asking appropriate questions."

S69: "I think that effectively using AI like ChatGPT can speed up learning, so it should be actively used."

Comments from the 'Others' category unearthed rich insights into students' comprehension of ChatGPT, its technical foundation, diverse functionalities, and potential applications in their academic and personal lives. Beyond its functional merits, students highlighted ChatGPT's capacity to distill complex information into digestible formats, tailoring its outputs to the learner's stage. Concurrently, some raised concerns about the ethical implications, academic honesty, and data privacy associated with its use.

Comments from the 'Others' category:

S29: "I'm aware that regulations are starting to be imposed on things like university theses."

S32: "I think it should be used only to a degree, as I'm afraid that excessive use might lead to plagiarism."

S84: "There may be instances of personal data leakage. It can summarize books for me."

3.4 Students' Expectation to the Course (Q.5)

Q5, the final question, invited feedback and suggestions on course management and English learning. It was optional, yet 78 students provided comments. Post filtration of generic responses like 'nothing special,' manual coding was performed. Figure 7 encapsulates the results.

A considerable number of students outlined specific objectives they hoped to achieve through the course. Their aspirations predominantly leaned toward enhancing specific English skills. It's worth noting that these targets might be influenced by the course's objectives. Surprisingly, extrinsic motivators like grades and course credit fulfillment were infrequently mentioned. The 'Class Management' category revealed diverse instructional preferences, suggesting an array of unique learner needs and expectations.

4. Impact on the Course Design

4.1 Students' AI Readiness for the Courses

The pre-course survey revealed that approximately 40% of the students consistently use AI editing tools. Furthermore, 80% had heard of ChatGPT, with about 20% utilizing the tool for various purposes, including as a learning aid. This data was surprising, given the somewhat ambiguous attitudes and evaluations from Japanese society and academia toward these tools in April 2023 (MEXT, 2023). The students had already engaged with them, recognizing their utility, and implemented them without any clear guidance or validated research on their effectiveness in learning.

In the absence of explicit governmental policies on the educational use of generative AI, the challenge was deciding whether to integrate them into the courses. There was a foreseeable risk: students might submit essays heavily reliant on

Figure 7
Themes of Students' Wishes to the Course (Q5)

Categories	Themes	Frequency 1	Frequency 2
Aspirations/Feelings	Anxiety/Weaknesses	10	
	Determination/Will to Try Hard	6	
	Excitement/Positive Anticipation	4	23
	Desire to Like/Enjoy English	3	
Purposes of English Learning	Future Career Aspirations	2	
	Graduate School Preparation	2	
	TOEIC Test Preparation	2	
	TOEFL Test Preparation	1	
	Improvement in Listening Skills	8	
	Improvement in Speaking Skills	9	
	Improvement in Writing Skills	8	66
	Improvement in Reading Skills	6	
	Improvement in All Four/Overall Skills	17	
	Grammar	3	
	Engagement Opportunities	4	
	Completion of English Course Credits	2	
	Aspiring for Good Grades	2	
Class Management	Instructional Preferences	8	8
Others	Others	7	7

these AI tools, making it challenging for instructors to distinguish original student work. Given the AI's ability to provide instant, tailored feedback, it offers advantages that human instructors can't match when dealing with a large number of students. Hence, it became imperative to provide clear guidelines on AI tool usage, ensuring optimal learning outcomes and avoiding potential pitfalls such as over-reliance.

It's essential to note that this survey specifically aimed to adapt the course design based on the students' AI readiness. The respondents were limited in number and directly related to the courses under review. However, the general sentiment among the surveyed students was in favor of incorporating AI tools into the course design, paving the way for the upcoming course revisions discussed in the next section.

4.2 General Course Design Parameters

A conducive ICT environment is crucial for the success of a tech-integrated course. The following design parameters were already in place, ensuring a technological infrastructure that supported the revised course design: (a) a stable internet connection within the school and its vicinity; (b) a BYOD policy ensuring that all students have personal laptops; (c) Moodle is available to all students for course management; (d) classrooms equipped with essential audio-visual equipment and

screens; and (e) given the institution's focus on science, students generally exhibit a higher aptitude towards technology.

The AI tools will primarily be incorporated into the English writing assignments, a common component across the five courses in this scheme. The surveyed courses aim to nurture academic English skills ranging from CEFR A2-C1, foster international awareness, and cultivate global perspectives. The courses differ based on their primary skill focus, with tracks dedicated to listening/speaking and reading/writing, even as they incorporate all four skill sets. All courses utilize assigned textbooks from Cambridge University Press (2019), which are organized around structured, multimedia-enriched activities spanning various cultures and topics. Adhering to school policy, all language courses are face-to-face, held weekly for 15 sessions each semester.

While the course structure is instructor-dependent, all courses in this study adopt a blended approach, merging online components with traditional classroom sessions, using Moodle as the primary platform. For the speaking-focused courses (CEFR B2 and C1), an AI voice recognition tool, otter.ai, has been introduced for audio recording and speaking practice. The integration of AI editing tools and ChatGPT is detailed in a later section. Post-course surveys will measure changes in students' perceptions and confidence in their skills, determining the revised design's efficacy. Additionally, two timed writing tests, simulating the TOEFL essay format, will assess skill progression and any unintended consequences of the adopted methodologies.

Across all courses, students will submit three major assignments: mini-speech scripts for speaking classes and short academic essays for writing classes, in addition to regular smaller tasks. The instructor will provide both numeric scores and written feedback for each assignment. Course grades will be determined based on overall performance, with no formal exams. Excluding tests and grades, fellow students can access and review submitted assignments. This approach aims not only to provide an audience for student work but also to promote transparency in course management and facilitate peer learning. Depending on the activity, Moodle's forum feature facilitates individual, paired, or grouped submissions.

4.3 AI Editing and ChatGPT in the Course Design

The use of AI tools is particularly envisaged during the assignment preparation phase where students must draft written content in English. All courses necessitate three assignments per semester: a short speech script (paired with an audio presentation) for listening and speaking courses, and an academic essay for reading and writing courses. Assignment length varies based on the student's English proficiency stage. For instance, basic-level assignments might be a single paragraph of 150 words, while advanced assignments could range from 500 to 800 words. In terms of audio recordings, these would last between three to eight minutes.

Two tiers of AI tools are planned for implementation. The first tier involves AI editing tools, targeted at correcting mechanical aspects such as spelling, basic grammar, and vocabulary. The second tier encompasses generative AI, focused on content-level advice, such as essay structure, cohesion, relevancy, and potential final editing or proofreading.

The combined use of Grammarly and ChatGPT was chosen for various reasons, the primary ones being their free version availability and browser-based compatibility. As of this study, ChatGPT stood out as the sole advanced linguistic generative AI tool accessible for free. Additionally, the survey

highlighted Grammarly as the most popular AI editing tool, used by 27 students or 22.5% (as shown in Figure 3), surpassing other alternatives like Writefull and ProWritingAid.

Using AI tools for speech script preparation is especially advantageous. It ensures that students don't consistently practice with scripts that might contain significant linguistic errors, potentially leading to the entrenchment of these mistakes.

The suggested procedure for using these AI tools is to: (a) draft an initial version to the best of one's ability; (b) utilize the AI editing tool for mechanical corrections; (c) seek advice from ChatGPT on content and organization. Conduct a final edit if required; and (d) self-check the final draft before submission.

Interactions with the generative AI are recommended to be predominantly in English to leverage the AI's linguistic abilities. However, depending on the students' readiness and course goals, mixing in their native language can also be an option, considering ChatGPT's multilingual capabilities.

For each assignment's final submission, students must provide five items: the first draft, the final draft, a reflection paper, screenshots of AI tool usage, and (for speeches) audio recording. While these are the recommended steps, students are encouraged to add further revisions as they see fit. The reflection report requires them to compare their initial draft with the final one, highlighting the learning achieved from the AI-assisted process. This reflective approach aligns with educational theories of discovery learning and the concept of 'noticing' (Schmidt, 1990).

The pre-/post-writing tests in the course design were discussed in tandem with the teacher's intention to gauge the impact of AI tool intervention. The course design, framed as an experimental setup, resonated well with the science-major students familiar with scientific experiments.

4.4 Other Considerations

For the AI editing tools and ChatGPT integration, additional materials and guidance are planned. First, an 'AI Tool Account Creation' guide covers the step-by-step process of setting up accounts for Grammarly, ChatGPT, and Google Documents (which can integrate with Grammarly's Microsoft Word plugin). The free versions of these tools are deemed adequate for course requirements. Second, an 'Engaging with ChatGPT' guide provides potential interaction scenarios, from initial draft consultations to final edit requests, distinguishing between prompts for oral or written English. Third,

Figure 8

Example Slide Illustrating the Assignment Creation Process

First Draft (By Yourself) → AI Editing (Edit 1) → ChatGPT (Edit 2) → Final Draft
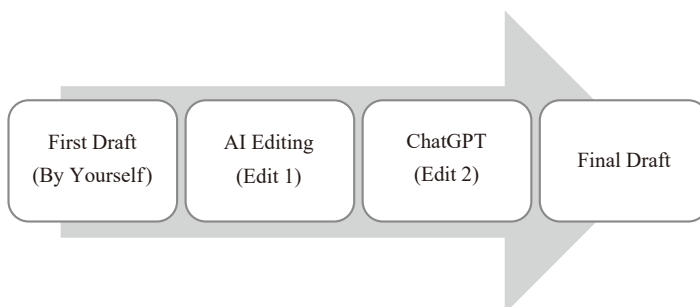
the decision to integrate AI tools across all proficiency levels is due to the unpredictable outcomes of AI tool implementation. A uniform approach avoids any biases that might skew the results. Fourth, course development will be continuously monitored. If adverse effects arise, course designs will be adjusted, including a potential revert to traditional teaching methods. Adhering to the ADDEI model, the revised course was launched in the Spring semester of 2023 without significant issues.

Finally, as of Fall 2023, while this paper is undergoing revisions, there have been significant advancements in generative AI technology, particularly in ChatGPT. A groundbreaking enhancement now enables voice-based, quasi-simultaneous communication with ChatGPT (Reuters, 2023). This opens new avenues for integrating the AI into speaking and listening-focused language classes. The ability to interact with ChatGPT through voice transforms it from a text-based tool into a dynamic, interactive platform, offering multi-lingual benefits for pronunciation, listening comprehension, and conversational practice. Following this, the introduction of 'MyGPTs' marks another notable evolution, allowing for the creation of custom versions of GPT (OpenAI, 2023b). Educators can now tailor the AI to specific needs, such as combining a 'Writing Coach' feature with classic GPT functionalities, without requiring extensive programming knowledge. This development is a substantial leap forward in making advanced AI tools more customizable and shareable in educational settings. These recent developments significantly expand ChatGPT's applicability in the realm of language education, aligning with the evolving needs of modern classrooms and offering innovative ways to enhance teaching and learning experiences.

5. Limitations

This study was conceived during the period from February to April 2023, a mere quarter-year after the emergence of generative AI technologies. While this rapid response to a nascent technology allowed for a timely exploration of its implications in English education, it also posed inherent limitations. The literature review for this paper was confined to research studies published from December 2022 to July 2023. While this timeframe captured the initial reactions and studies pertaining to generative AI in education, it might not encapsulate the full breadth of academic responses that could have emerged shortly after. Given the fast-paced evolution of AI, especially generative models, the landscape of technology and its educational implications can change swiftly. The findings and observations made in this study are based on the state of technology and its adoption as of mid-2023, and future advancements may provide new insights or shift the dynamics observed in this research. The research studies included in the literature review were primarily sourced from Scopus, potentially omitting valuable insights or findings present in other databases or journals not indexed therein.

6. Conclusion

The integration of advanced AI language models into our pedagogical practices is an ongoing and rapid evolution. As these generative AI models continue to enhance their voice and textual dialogue capabilities, they promise to further bolster both daily life and learning. Crucially, the outputs generated by ChatGPT reflect our inputs—it serves as a mirror to our inquiries. As educators, it's vital to recognize the growing appeal of such technologies among students. They represent not just an additional tool, but potentially a new

paradigm in teaching and learning. To remain relevant, educators must embrace and adapt to this emerging technology; otherwise, traditional language teaching roles might become obsolete, overshadowed by these advanced AI systems.

7. Acknowledgement

The successful completion of this research was made possible through the unwavering cooperation of my students at the Tokyo University of Science, my primary institution. I also extend my gratitude to the Institute of Educational Research and Service of International Christian University, where I am affiliated as a research fellow, for their support. Lastly, AI editing tools and ChatGPT played an indispensable role in refining and enhancing this paper. I look forward to sharing the methodologies behind this work with my students, hoping it aids their growth in English proficiency, career trajectories, and broader life experiences.

8. Appendix

Pre-coursal Survey on AI Editing Tools and ChatGPT (April 2023)

Request for Cooperation

This survey is intended to help us better understand how you are learning English.

The content of the questions is about learning tools for learning English.

Your responses are kept anonymous, there are 10 questions, and the time required to answer is 5 minutes.

Your responses will not affect your course grades.

Responses may be aggregated, quantified, and published in conference presentations, papers, etc.

Thank you for your cooperation!

Course being Taken

Please select the courses you are currently enrolled in. (Multiple answers possible)

Advanced English 1
Listening and Speaking 3
Reading and Writing 1 (Thursday 4th Period)
Reading and Writing 3 (Thursday 5th Period)
Reading and Writing 1 (Thursday 6th period)
Language and Culture 1

Regarding English

Q1.1 Out of the four English language skills listed below, which do you enjoy the most? (Multiple answers possible)

Reading
Writing
Listening
Speaking
None of the above

Q1.2 Which of the following English language skills do you feel you excel at? (Multiple answers possible)

Reading
Writing
Listening
Speaking
None of the above

About AI English Editing Software

Q2 Regarding English editing software using AI (grammar and spelling check, etc.), please choose the one that best describe your experience so far.

I use it frequently.
I use it occasionally.
I've used it once or twice.
I've heard of it but have never used it.
I've never heard of it (or don't know it).

AI English Editing Software (continued)

Q2.1 Please choose the AI English editing software that you have used. (Multiple answers possible)

Microsoft Word's English proofreading feature
Writefull
Grammarly

Pro Writing Aid

Other (fill in)

About ChatGPT

Q3.1 Regarding ChatGPT, please choose the one that best describe your experience so far.

I use it frequently.

I use it occasionally.

I've used it once or twice.

I've heard of it but have never used it.

I've never heard of it (or don't know it).

Q3.2 Feel free to write in the below box what you know about ChatGPT at the moment, what you think, how to use it, etc.

ChatGPT (continued)

Q3.3 For what purposes have you used ChatGPT? (Select all that apply.)

English proofreading (grammar/spell check)

Content consultation and advice

Composition and structuring of sentences

Writing style check

Reading support

Listening support

Speaking assistance

Other (fill in)

Respondent profile

Q4.1 Please choose the one that best describes your major field of study.

Mathematics (including Applied Mathematics)

Physics (including Applied Physics)

Chemistry (including Applied Chemistry)

Architecture

Management (including Business)

Other (fill in)

Q4.2 Please choose one that applies to your age.

18

19

20

21

22

23

24

25 years old or older (fill in the number if you like)

I chose not to answer this question.

Q4.3 Please select one answer that applies to your gender.

Male

Female

Other

I chose not to answer this question.

Others

Q5 Please feel free to write your requests and opinions about this course.

End of Survey

We thank you for your time spent taking this survey.

Your response has been recorded.

References

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, *33*, 1877-1901. https://doi.org/10.48550/arXiv.2005.14165

Business of Apps. (2023, May). ChatGPT Revenue and Usage Statistics (2023). https://www.businessofapps.com/data/chatgpt-statistics/

Cambridge University Press. (2019). *Unlock* (Levels 2-5) (2nd ed.). Cambridge University Press.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, *13*(1), 3-21.

Council of Europe. (n.d.). The Common European Framework of Reference for Language (CEFR) Levels. https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications.

Ehara, Y. (2023). Innovative software to efficiently learn English through extensive reading and personalized vocabulary acquisition. *Communications in Computer and Information Science*, 1831, 187-192. https://doi.org/10.1007/

978-3-031-36336-8_28

Elsevier. (n.d.). Scopus. https://www.elsevier.com/solutions/scopus

Felten, E., Raj, M., & Seamans, R. (2023). How will language modelers like ChatGPT affect occupations and industries? https://arxiv.org/ftp/arxiv/papers/2303/2303.01157

Gagne, R. M., Wager, W. W., Golas, K. C., Keller, J. M., & Russell, J. D. (2005). *Principles of instructional design* (5th ed.). Thomson Wadsworth.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*. https://doi.org/10.1177/00336882231162868

Kovačević, D. (2023). Use of ChatGPT in ESP teaching process. 2023 22nd International Symposium INFOTEH-JAHORINA, INFOTEH 2023. https://doi.org/10.1109/INFOTEH57020.2023.10094133

Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2023). Daigaku-Kousen ni Okeru Seisei AI no Kyougaku-men no Toriatsukai ni Tsuite (Shuuchi) [Handling of Generative AI in the Educational Aspect at Universities and Colleges of Technology (Notification)]. https://www.mext.go.jp/content/20230714-mxt_senmon01-000030762_1.pdf

Mohamed, A. M. (2023). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: Perceptions of EFL faculty members. *Education and Information Technologies*. https://doi.org/10.1007/s10639-023-11917-z

Tokyo University of Science (TUS). (2023). Number of undergraduate students. https://www.tus.ac.jp/about/data/number_students/faculty/ [accessed July 25, 2023]

OpenAI. (2022, November 30). Announcement. https://openai.com/blog/chatgpt

OpenAI. (2023a, March 27). GPT-Technical Report. https://arxiv.org/pdf/2303.08774

OpenAI. (2023b, November 6). Introducing GPTs. https://openai.com/blog/introducing-gpts

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI. https://d4mucfpksywv.cloudfront.net/better-language-models/language-models

Reuters. (2023, September 25). OpenAI's ChatGPT will 'see, hear and speak' in major update. https://www.reuters.com/technology/openais-chatgpt-will-see-hear-speak-major-update-2023-09-25

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129-158. https://doi.org/10.1093/applin/11.2.129

World Economic Forum. (2023). Fourth industrial revolution. https://www.weforum.org/focus/fourth-industrial-revolution

Young, J. C., & Shishido, M. (2023). Investigating OpenAI's ChatGPT potentials in generating chatbot's dialogue for English as a Foreign Language learning. *International Journal of Advanced Computer Science and Applications*, 14(6), 65-72. https://doi.org/10.14569/IJACSA.2023.0140607