Assessing AI-based Summarizing and Paraphrasing Tools for a CLIL Intercultural Communication Academic Writing Class

Barry KAVANAGH
Tohoku University, Japan

Abstract: This paper aimed to examine the benefits and drawbacks of AI-based summarizing and paraphrasing tools that were employed in a CLIL academic writing course on intercultural communication at a national Japanese university. Students summarized an internet article of their choice that had to include indirect citations by paraphrasing the author's ideas. First, they did this without the support of AI-based tools. On their second and third attempt, they used the AI-based summarizing and paraphrasing tools Quillbot and SpinBot respectively. A one-way ANOVA was conducted to evaluate the three summaries for grammar, readability and plagiarism. Results showed that although the AI-based tools can improve grammar, they have limited capability in improving readability and preventing plagiarism. A survey on student perceptions of the tools also highlighted the drawbacks of the software and that dependence on such tools to produce academic writing content needs to be treated with caution.

Keywords: AI-based tools, summarizing and paraphrasing, academic writing, plagiarism

1.　　Introduction

The research field of artificial intelligence in education (AIEd) aims to focus on the development and subsequent improvement in how computer software can improve teaching and learning. Within the field of teaching English as a second language, artificial intelligence (AI) has been an integral part of the evolution of how computer technology has been used in the language learning classroom. This led to the advent of Computer Assisted Language Learning (CALL) in the 1960s and in the 1970s to Intelligent computer-assisted language learning (ICALL), a multidisciplinary area of research that combines natural language processing (NLP), intelligent tutoring system (ITS), second language acquisition (SLA), and foreign language teaching and learning (FLTL) (Tafazoli et al., 2019).

In a 2018 Horizon report, experts suggested that AI in education would grow by 43% from 2018-2022 (Educause, 2018). This growth has been accelerated with the onset of the COVID-19 pandemic and has led to research outlining the new normal within education post-COVID-19 (See Sintema, 2020 on digitalized virtual classroom; Basilaia & Kvavadze, 2020 on online education; Naciri et al., 2020 on mobile learning; Mulenga & Marbán, 2020 on digital learning).

The COVID-19 pandemic saw a huge shift in how education was delivered to students with universities adopting online platforms such as Google Meet and Zoom to conduct classes. This led many teachers, who were previously accustomed to traditional face-to-face classrooms, to innovate and use online tools and technology that on the whole was a new experience for them. As students returned to the classroom many universities have now adopted a hybrid system that has combined both in-person and online classes. Universities have also implemented a 'Bring Your Own Device' (BYOD) policy whereby students are required to bring and use a mobile device in their classes. This has led to many teachers using online tools to aid them in their teaching. The use of artificial intelligence in education can divide opinions within the research literature with some citing the

benefits it that has to offer whilst others have expressed fears and doubts about the reliability of the technology.

This research aims to examine the benefits and drawbacks of AI-based summarizing and paraphrasing tools that were employed in a CLIL academic writing course on intercultural communication at a national Japanese university. The main issues it will try to address are whether or not such tools can improve a student's academic writing (grammar, structure, style), help them avoid plagiarism, and provide them with the skills to paraphrase sources through indirect citations.

2.   AI Usage in Language Education: Praise and Points of Contention

AI was first coined by John McCarthy in 1955 and since then AI research has been conducted across many academic disciplines and produced a considerable body of literature
(e.g., Burleson & Lewis, 2016; Kaplan & Haenlein, 2019).
Nilsson (2011) has suggested that AI is about making intelligent machines based on the model of human thinking and the intelligence of humans. Wartman & Combs, 2018 state that AI is generally defined as the ability of machines or computers to replicate human thinking and to think and act like humans through imitation. In language learning and teaching, it is suggested that AI can duplicate the behavior of the teacher (Bailin, 1987; Matthews, 1993).

However, within the field of English as a Second Language (ESL), opinion is polarized on how effective it is in helping learners become more proficient in the language. Many teachers have not fully embraced the technology and do not wish to use it in their classrooms (Prensky, 2008; Kaban and Ergul, 2020; Istenic et al., 2021). Researchers and teachers who have found it a useful aid have cited that it improves a student's grammar and can provide feedback (Bailin, 1987). More recent studies have shown that the integration of AI technology in foreign language education can provide flexible, interactive and student-centered learning opportunities (El Shazly, 2020) and that AI technology can help students meet their L2 goals and improve reading ability (Bailey et al., 2021). Yin et al (2021) investigated the impact of a chatbot-based micro-learning system and found that it increased students' motivation.

Salaberry (1996) suggests that many teachers have expressed disappointment with previous technological 'revolutions' which can lead to them being less receptive to the pedagogical uses of this new medium. Gallacher et al. (2018) examined student perception of a chatbot used in the English language learning classroom and found that students viewed it as a novelty rather than a language learning tool. They suggest that language teachers need to be more critical of AI technology and be cautious when incorporating it into the language learning classroom.

3.   The Use of Summarizing and Paraphrasing Tools in Academic Writing

Good academic writing can reflect the author's skill in paraphrasing a source to demonstrate that they have understood the nuance of what they have read and also appropriately cite the source when making indirect quotations that rely on the author using their own words to express the ideas written in the source material (Keck, 2006, 2014; Shi, 2012).

There is now a prevalence of easy-to-access summarizing and paraphrasing tools, both paid and free versions, which may or may not have limited functionality. Niño (2009) suggests that such software is questionable from an educational standpoint and Rogerson & McCarthy (2017) stipulate that these Internet-based paraphrasing tools are text-processing applications and are associated with the same approaches used for machine translation (MT).

These tools aim to help students summarize or paraphrase material when writing essays by offering alternative sentences, expressions, grammar, and vocabulary so that it is sufficiently different from the source material but retains the same meaning and nuance.

Studies on the usage of summarizing and paraphrasing tools within the English language classroom are absent in the research literature, especially within classes that focus on academic writing. This paper aims to address this gap in the literature by examining how the AI-based summarizing and paraphrasing tools 'Quillbot' and 'SpinBot' are perceived by students who used them in a CLIL intercultural communication academic writing class. This paper aims to address the following research questions:

1. Do AI-based online tools improve students' grammar when writing summaries and paraphrasing content?
2. Do AI-based online tools improve the style (readability, structure) of a student's writing?
3. What do students find difficult in summarizing and paraphrasing when writing with and without the support of AI tools?
4. What are the pros and cons of using AI-based software for academic writing according to the students who used them?
5. Can these AI-based tools help students avoid plagiarism?

4.   Data

4.1   The Course

The assessment of AI-based summarizing software and how students perceived it, was conducted in an 'Academic Writing' course at a national Japanese university for second-year students. The 33 students in the class were engineering majors. Based on the entrance examination test scores of the university the students were classified as being advanced. A textbook called *Pathways to Academic English* (2021) written by a team of teachers (author included) was used for the course. The textbook covers all the course content that the students learn in both the first and second years of their English studies. Two chapters are dedicated to this 'Academic Writing' course. The chapters act as a guide for students to learn how to write a 5-paragraph essay based on an introduction, body, and conclusion (IBC) format and how to cite and write references. The objective of the course was for students to acquire the ability to write an academic essay.

To teach the course the author incorporated a CLIL approach whereby students learn a subject through the medium of a foreign language enabling them to both learn the content of that subject and the foreign language simultaneously**.** Coyle et. al, (2010) define CLIL as "a dual-focused educational approach in which an additional language is used for the learning and teaching of both content and language" (p. 9). The content of this class focused on units related to intercultural communication. The focus, therefore, was on a language-driven 'soft' CLIL approach whereby content is used to learn the L2, and language learning is the priority (Ohmori, 2014). Based on the chapters in the *Pathways to Academic English* textbook, students learned the following:

- How to write a good essay introduction (hook, background information, thesis statement)
- How to construct the main body of their essay
- How to write an effective conclusion
- The basics of citing (indirect and direct quotations – use of reporting verbs – paraphrasing – summarizing sources – avoiding plagiarism).
- Referencing – how to write references based on particular styles such as APA.

As the content of the course was on intercultural communication the students had to write an academic essay based on the course content and were free to choose their own intercultural communication topic or theme for their essay. This could be content covered in the course or something related that was of particular interest to the student.

4.2   The AI-based Summarizing and Paraphrasing Tools Used

There is a huge variety of summarizing and paraphrasing tools available, both paid and unpaid. For this study, Quillbot and Spinbot were chosen based on their reputation (they both often appear in the best-ranked summarizing and paraphrasing tools lists).

Quillbot
   Quillbot is an online paraphrasing and summarizing tool. The website states that "QuillBot's paraphrasing tool helps millions of people rewrite and enhance any sentence, paragraph, or article using state-of-the-art AI." As shown in figure 1 Quillbot offers many functions as reflected in the column on the left-hand side.
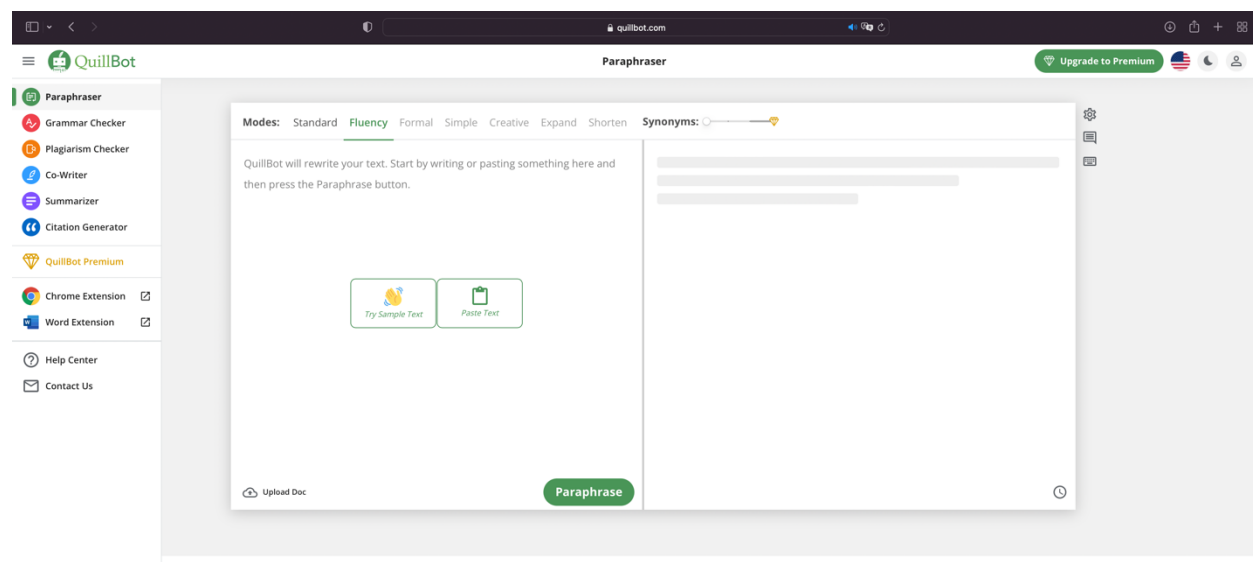

Figure 1. Screenshot of the Quillbot Online Summarizing/Paraphrasing Tool

   The user pastes in their text into the left box and the results appear in the box to the right. The plagiarism function is only available for the paid version. The paraphraser, grammar checker, and summarizer can be used with the free version, but it is limited to 125 words per pasted text entry.

SpinBot
   SpinBot is marketed as an 'Intelligent, Free Text Rewriting Tool' where you can rewrite (summarize text) through the text spinner option or paraphrase your inputted text through its paraphrasing tool. The website states "SpinBot is a free, automatic article spinner that will rewrite human-readable text into additional, intelligent, readable text. Similarly, if you need a paraphrasing tool, SpinBot will get the job done for you for that purpose as well."

SpinBot is free to use, although there is a paid version that is advertisement free. You can input up to 10.000 words at a time. Below is a screenshot of SpinBot. You simply paste your text into the left box, and you can see the results it produces in the box to the right. The same applies to the paraphrasing tool.
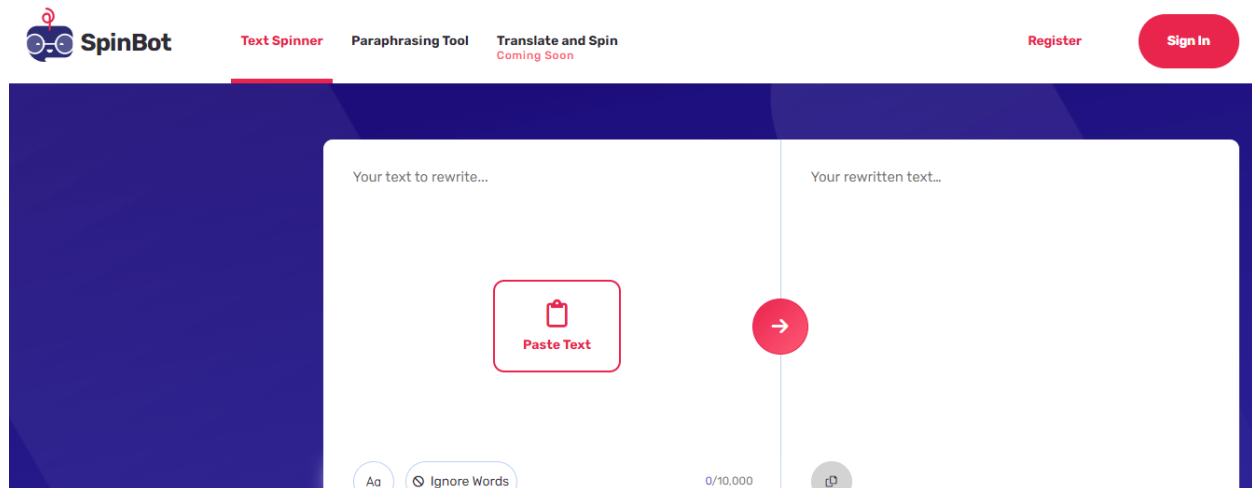

Figure 2. Screenshot of the SpinBot Online Summarizing/Paraphrasing Tool

5.   Methodology (Research Procedure and Classroom Practice)

5.1   Assignment

Students had to find an article on the internet based on the topic of the academic essay they decided to write their paper on. Students decided on their essay title by week 7 of the 15-week course which allowed the students the remaining 8 weeks to do some research on their essay and write it up. The paper had to be at least 3 pages in length with references. This assignment was designed to get the students to do online research by looking at English websites (a list of good websites was also provided to the students) and through reading them deciding on what might be useful to use for their essay. Students were instructed to do the following:
- Please do an internet search and find an article that interests you and a topic you are thinking about writing your final essay on. The theme of the article you choose must be related to an intercultural communication topic similar to the themes we have done in class.
- The internet article must be at least 1.000 words in length and have an easily recognized author.
- Please consult the website listings provided in your handout which might inspire and help you with the assignment.
- Read the article you chose carefully and look up any unfamiliar words or expressions.
- In the next class, you will be placed into small groups and will orally summarize the content of the article that you chose.

5.2   The First Class: Summarizing and Paraphrasing the Internet Article Using Reporting Verbs and Indirect Quotations

For this assignment the students had to do the following:

1.   Students were placed into groups and gave an oral summary of their chosen article.
2.   Students were then told that they had to summarize what the article was about in 250-350 words (done in class through Google forms). Students were given 25 minutes to complete the task. In their summaries, they also had to include indirect quotations through the paraphrasing of what the source (original author) stated. This is the reason that the articles they used required an easily recognizable author. Before this class, students were taught that to make an indirect quotation they needed to paraphrase the original author's idea using their own words and include a reporting verb and the author's name when making the citation. This was reviewed again in this class along with the basics of how to write up a good summary with paraphrased content for indirect quotations.
     When writing their summaries students were told to:
•   Understand the main idea / supporting details of the source material.
•   Use your OWN words to summarize the main point of the passage and use at least one indirect citation in your summary.
•   Paraphrase with reporting verbs on what the author thinks or says.
•   Use reporting verbs (suggest, states, according to, etc.) to help you cite and paraphrase the author's words.
•   Change the parts of speech of some words.
•   Consider changing the writing from passive to active voice or vice versa.
•   Look for ways to utilize reduction – for example, changing a clause to a phrase.
•   Replace some words in the original text with synonyms.
•   You are allowed to use dictionaries and thesaurus for this activity.
     After students completed the summary task, all students were given a Google form survey to complete based on their experience of doing this assignment.

5.3   The Second Class: Summarizing and Paraphrasing an Internet Article Using Reporting Verbs and Indirect Quotations with AI Software Support

To assess whether AI software would be beneficial and create better summaries and paraphrasing in comparison to their work without AI support, the students were asked to summarize and paraphrase the same chosen article from last week, but this time use AI-based online tools. First, students were introduced to the Quillbot software (free version). The 125-word limitation was explained, and students were given a demonstration on how to use the software. Students were asked to use their judgment of the summary produced by the software and change it accordingly. They did not have to accept the output of the software if they found that it produced strange or awkwardly phrased sentences or expressions. Quillbot has a 'rephrase' option where you can change parts or whole sentences with choices given from drop-down menus.

As the students had never used or heard of such software before they were given a demonstration lesson on how to use the tools and the functions that were available to them with these free versions. The limitations of the free versions were also explained. Students were then given a piece of text through Google Classroom on an intercultural communication topic of just

over 1.000-words that was studied and used in a previous class and was not new to the students. Students were then instructed to paste the text into Quillbot and then SpinBot and get familiar with the functions they provide. Further explanations and help were given to the students where necessary.

Students were then told that they would write another summary of the article they chose last week but this time with the help of the two AI-based software platforms. Students were given the same instructions as in the previous class. The only difference is that they pasted their chosen articles (in parts for Quillbot due to the125 word limitation) into the AI tools and worked on their summaries with the software support. Students could refer to the summaries they did last week if they wished. Students were given 25 minutes to write each summary. One using Quillbot and the other using SpinBot. Student summaries were again written in Google forms.

After students completed the summary task, all students were given a Google form survey to complete based on their experience of doing this task using AI-based online tools and their opinions on the results it produced.

5.4   Evaluation of Student Summaries – Pro Writing Aid

For this study and to continue using AI as the core of this study, student summaries were evaluated using the 'Pro Writing Aid' software. This allowed for subjectivity and also to evaluate the benefit of using such software. (Actual grading and marking of the summaries were done later by the teacher). The website claims that "Pro-Writing Aid is the only platform that offers world-class grammar and style checking combined with more in-depth reports to help you strengthen your writing." It can also check writing based on genres such as formal academic writing and informal writing.

To evaluate student summaries based on an objective AI-based software all 3 of the student summaries were run through the software and their score for 'grammar' and 'style' were recorded for each summary and then compared for statistical differences using a one-way ANOVA one test and an alpha level of .05 was set. The grammar score is the result given based on a scan of the inputted documents with misspelled words, missing punctuation, their structure, and tenses. Pro-writing aid provides a style report on the inputted text with suggestions on how to improve the document. It suggests ways to make your writing more readable, and less complex and offers suggestions on hidden, passive verbs, long subordinate clauses, and a passive index. Pro-Writing Aid suggests that it is one of the most popular and comprehensive reports that the software offers.

To evaluate student summaries based on an objective AI-based software all 3 of the student summaries (1. No AI support, 2. Support with Quillbot, and 3. Support with SpinBot) were run through the software and their score for 'grammar' and 'style' were recorded for each summary and then compared for statistical differences using a one-way ANOVA test.

5.5   Checking for Plagiarism

Students' summaries were checked for plagiarism using the 'Plagscan' software to see if there were any differences in terms of similarity to the original article between the summaries with and without software support. It must be stressed that Plagscan, like most plagiarism checkers, provides a summary of matching or similar text of the submitted work comparing it to a huge database of internet sources. It does not necessarily prove plagiarism but simply highlights where the text is similar to the source found in the database. The similarity scores provided by

Plagscan, therefore, show how much of the text matches online sources but does not explicitly suggest that this is the result of plagiarism. Therefore, it is up to the user of such software to determine which of the text warrants the accusation of plagiarism. Plagscan highlights the text in different colors to help the user in making such decisions.

After running the summaries one at a time through Plagscan, the similarity report shows sentences that are an exact match to the source material in red. Sentences that are paraphrased from the source were highlighted in blue and text that is a direct quotation was shown in green. Red sentences were only included in this analysis after going through the 96 summaries.

Legend text highlighting

Aaexact match
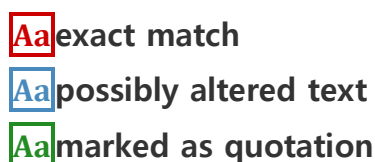Aapossibly altered text
Aamarked as quotation

Figure 3. Plagscan Text Highlighting Colors

6.   Results

32 of the 33 students completed both sets of summaries with and without AI support. The results are broken up into three main sections, student summary scores, plagiarism test scores, and after-summary completion survey results.

6.1   Student Summary Scores

Each of the three summaries, the summary without AI support, and the other two summaries which were written with the support of Quillbot and Spinbot were pasted into 'Pro Writing Aid' and a score for grammar and style for each of the 3 summaries was provided.
A one-way ANOVA was conducted to compare the Pro-Writing Aid evaluation of grammar and style scores of the three student summaries and the results are shown in tables 1 and 2.

Table 1. The Pro-writing Aid 'Grammar Scores'

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3040.75	2	1520.375	5.15703	.007517
Within Groups	27417.9063	93	294.8162		
Total	30458.6563	95			

The Pro-writing aid 'grammar scores' showed that there was a statistically significant difference in the grammar scores between at least two of the summary groupings $F_{(2, 93)} = 5.15$, $p < .05$. Post hoc comparisons using Tukey's HSD test indicated that the mean of the grammar score for

the student summaries with 'no support' (M=41.22, SD=17.72) was significantly different than the grammar scores for the summaries written with Quillbot support (M=54.03, SD=20.22). In addition, the mean of the grammar score for the Quillbot support summary (M=54.03, SD=20.22) was significantly different than the mean of the grammar scores for the SpinBot support summaries (M=43.21, SD=12.70). However, there was no difference in the mean scores of the summaries with no AI support (M=41.22, SD=17.72) and the summaries written with the support of Spinbot (M=43.21, SD=12.70).

These results suggest that summaries written by students with the Quillbot software led to better grammar scores based on the results within the Pro-writing aid platform. Grammar scores of the summaries without AI support and written with the aid of SpinBot showed no differences which suggests that SpinBots contribution to improving the student's grammar within their summaries was minimal and not significant. Quillbot, however, improved students' grammar scores significantly when compared to the no AI support and Spinbot summary groupings. These findings suggest that only the Quillbot tool could make an impact on students' grammar scores.

Table 2. The Pro-writing Aid 'Style Scores'

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2292.1458	2	1146.0729	6.03497	.003433
Within Groups	17661.1875	93	189.9052		
Total		95			

Results of the 'style' scores showed that there was a statistically significant difference in the grammar scores between two of the summary groupings F (2, 93) = 6.03, p < .05.
Tukey's HSD test indicated that the mean of the style score for the student summaries with 'no support' (M=58.37, SD=14.05) was significantly different than the style scores for the summaries written with SpinBot support (M=46.40, SD=11.52). There was no statistical difference between the SpinBot and Quillbot scores or within any other summary group pairings.
Looking at the mean scores of the 3 summary groups, both the Quillbolt mean score (52.46), and the SpinBot mean score (46,40) were lower than the summary mean scores without AI support (M=58.37). This suggests that both sets of AI support made the style scores worse and did not have a positive impact on the style of the summaries. This would suggest that the readability of these summaries deteriorated with the use of these AI-based tools according to these Pro-Writing Aid evaluations.

6.2   Plagiarism Test Scores

The result of the one-way ANOVA showed that there were no statistically significant differences between the mean scores of the Plagscan results for the summaries written without AI support and with Quillbot or Spinbot support ($F(2,93) = 1.976$, $p = .1443$). Although, there were instances of the number of matching sentences decreasing when using Quillbot or Spinbot there were also cases where the opposite was true. Some of the student's original summaries without AI support had none or very few matching phrases or sentences but this increased when the same students wrote their summaries using Quilbot or Spinbot. Therefore, the question of whether or not AI paraphrasing software can help prevent plagiarism was inconclusive. This meant that the AI

software did not always seem to change or paraphrase enough of the content for it not to be flagged by the software as matching the source.

Plagscan scans documents and bases the results on the following similarity-level criteria:

0-1% The document is unlikely to contain any plagiarism from the internet or local databases.
1-5% A closer look at the document is recommended
5% and over: The document most likely contains plagiarism and an in-depth look at the document is required. This criteria, therefore, is just a recommendation for the user to look for plagiarism and the software provides the URL of the source material that the submitted work is similar with.

Below are the results of the student summaries and how they were rated by Plagscan.

Table 3. Percentage of Matching Sentences

	Summaries with no AI Support	Summaries with Quillbot Support	Summaries with Spinbot Support
Summaries with a score of 0-1	15	13	8
Summaries with a score of 1-5	1	2	0
Summaries with a score of 5-100	16	17	24

Half of the summaries with no AI support registered as having up to 5% similar red highlighted content to online sources, a similar result was found with summaries using Quillbot. The number dropped to a quarter (a total of 8 students) with the SpinBot software. AI support, therefore, did not reduce the number of matching sentences and they increased when using SpinBot. Students who scored from 5-100 averaged a score of 21 with no support, 17 with Quillbot support, and 22 with summaries written with Spinbot support.

6.3   After Summary Completion Survey Results

After completing the first summary assignment without the use of AI support, students were asked about what they found to be difficult with the task. The multiple choice question gave students choices that were based on the instructions given to students on what they had to do to write the summary and what they had learned about how to compose a good summary with indirect quotations.

What did you find difficult when summarizing the article you chose?

- General paraphrasing (Putting the article into your own words)
- Formulating and putting sentences togather
- Deciding and working out the main idea in an essay
- Writing the supporting ideas in the article I chose
- Changing the active sentences in passive voice
- Changing the vocabulary, word order or structure of the sentence
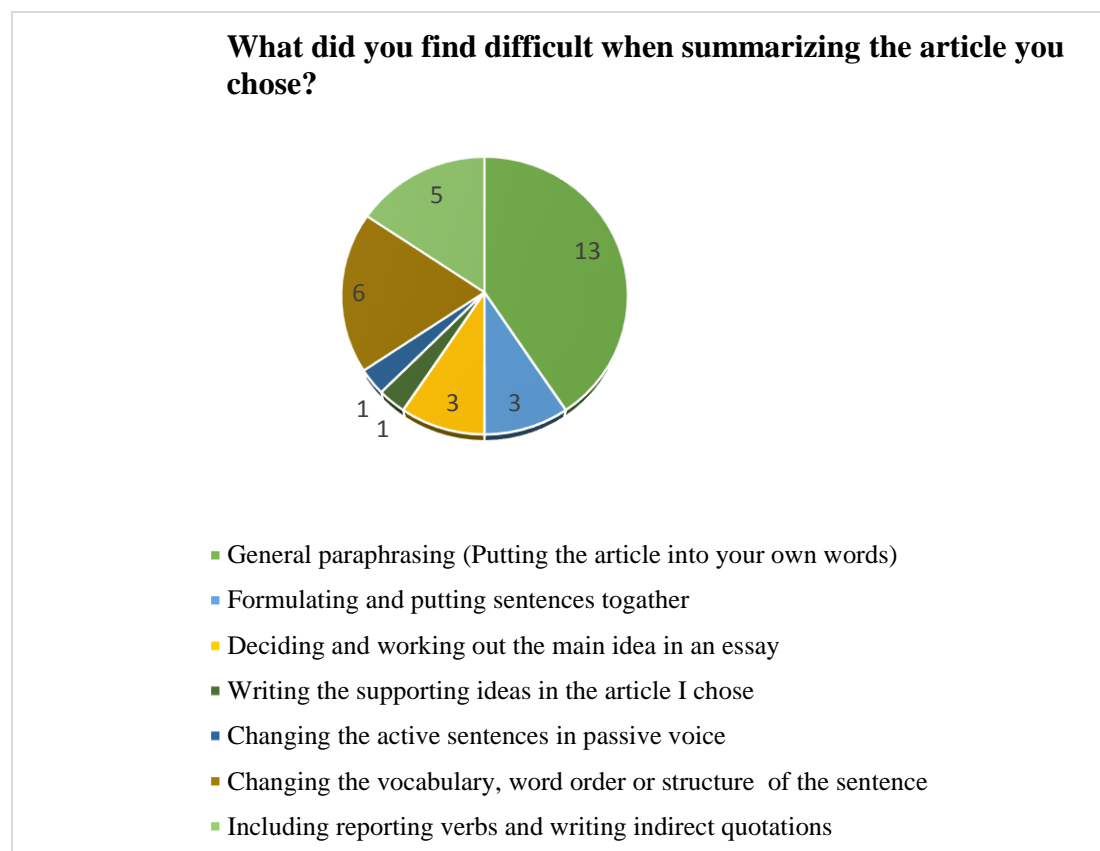- Including reporting verbs and writing indirect quotations

Figure 4. Student Reaction to the First Summary Assignment without AI Support (N=32)

Of the 32 respondents, results showed that 41% (13 students) found the general task of putting the article into their own words through paraphrasing was the most difficult part of writing the summary. This was followed by the specifics of paraphrasing such as changing the vocabulary and structure of the sentences in the article (19% of (6) students). 16% (5 students) stated that writing indirect quotations using reporting verbs was difficult. These results suggest that the general task of paraphrasing parts of their chosen article in their own words was the most difficult aspect of this task.

After completing the second-class summary assignment with the use of AI support through Quillbot and SpinBot, students were asked what they found to be difficult with the task whilst using the software.

What did you find difficult when summarizing the article you chose using AI based tools?

- ■ General paraphrasing (Putting the article into your own words)
- ■ Formulating and putting sentences togather
- ■ Deciding and working out the main idea in an essay
- ■ Writing the supporting ideas in the article I chose
- ■ Changing the active sentences in passive voice
- ■ Changing the vocabulary, word order or structure (using AI tools) of the sentence
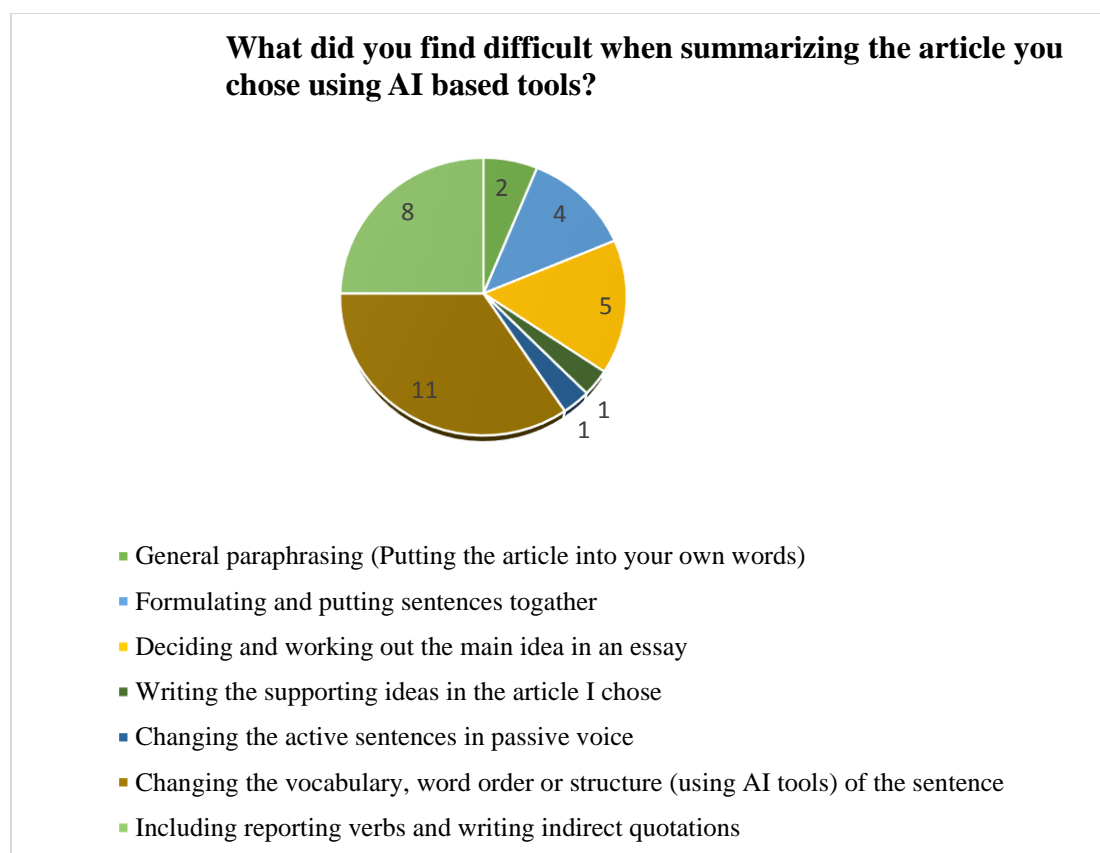- ■ Including reporting verbs and writing indirect quotations

Figure 5. Student Reaction to the Second Summary Assignment with AI Support (N=32)

Of the 32 respondents, results showed that only 2 students found that general paraphrasing was the most difficult. Students suggested that the software was generally good when helping students to paraphrase. This was perhaps expected as the sole purpose of the software is to help students summarize and paraphrase text. 11 students (34%) stated that changing the vocabulary, word order, and structure of the summary produced by the software was difficult. Students suggested that some sentences produced by the AI tools were too informal, unnatural, or awkward and they had to correct and edit them. They stated that the results were often unreliable and that it was time-consuming to correct them. These findings run parallel to the data in table 2 above that shows that style scores were not improved by the AI tools. Similar to the first summary assignment without AI support, some students suggested they had difficulty using reporting verbs and writing indirect quotations. This is perhaps not surprising as the software is not intended to assist in writing indirect quotations.

Students were asked to state their opinions on the pros and cons of the two AI applications. A summary of the main student comments (unaltered or corrected) are shown in Tables 5 and 6.

Table 5. Advantages and Disadvantages of Quillbot

Advantages	Disadvantages
Relatively accurate	It cannot summarize a text of more than 125 words.

The quality of the paraphrasing is better that Spinbot. The number of grammatical errors is not so big. You can see a lot of synonyms.	The limitation of words you can paste one time is only 125, so you should repeat copy and paste again and again.
Grammatical mistakes are relatively few	Sometimes, it chooses unsuitable words for the summary.
It can make sentences to suit the context. We do not have to edit much to make good sentences.	I think revised sentences may be a little informal.
Quillbot gives us a lot of choices of words if we want to change sentences.	It changes phrases too significantly.
Quillbot offers natural paraphrased words.	Sometimes it changes too much, so I can't tell what I want to tell.
Ability to vary the degree of correction.	If I made grammatical mistakes, the software gave strange expressions.
The software can both summarize and paraphrase sentences and choose the length of the summary.	The sentences are a little long.

Table 6. Advantages and Disadvantages of Spinbot

It can paraphrase vocabulary and keep the form of original sentences.	Spinbot tends to use difficult words and phrases, so they are difficult for us Japanese to emulate.
Spinbot tends to use difficult words and phrases, and many nouns, but they are desirable for formal writing like this essay.	The quality of the paraphrase is not so high.
Spinbot often proposes alternative words which look intelligent	It is too eager to change all the words that it often misunderstands the real intention of the texts. For instance, it paraphrased "what the world would be like" to "what the world resembles."
You can paste 10000 words	The sentence is not natural.
It can summarize a text, of more than 125 words.	There are many grammatical errors, especially prepositions
Can learn many kinds of expressions I rarely use.	The summaries we rewrite are quite different from what we write at first. We can say the new ones are not our summaries.
Being able to compare the sentences before and after paraphrasing	This application is not very good for paraphrasing.
The design is easy to understand.	Having to correct tenses and articles myself.
It can paraphrase vocabulary and keep the form of original sentences.	There are many advertisements.

Although students had reservations about the results that AI tools produce, predominantly the free SpinBot AI software, the majority of students (see figure 6 below) felt that these tools can be useful in writing essays in the future. Students commented that Quillbot was the better choice for this but that the 125-word limit was a drawback, along with the assertion that they cannot afford to buy the paid version which has full functionality.

Do you think this kind of software is useful for you to write essays?
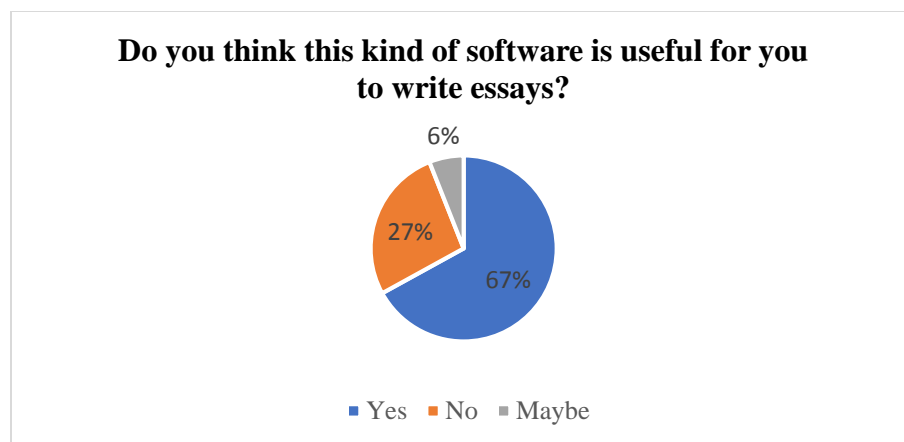
6%

27%

67%

■ Yes ■ No ■ Maybe

Figure 6. The Benefit of AI Support in Writing Essays (N=32)

Students were divided on whether or not AI tools of this type can successfully prevent plagiarism (See figure 7). 36% said 'yes', 22% said 'no' while the majority were on the fence and said 'maybe.' These opinions are also reflected in the findings which showed that there were no significant differences in the student summaries when they were scanned for matching sentences using Plagscan.

Do you think summarizing and paraphrasing software like Quillbot and SpinBot can prevent plagiarism?

42%   36%
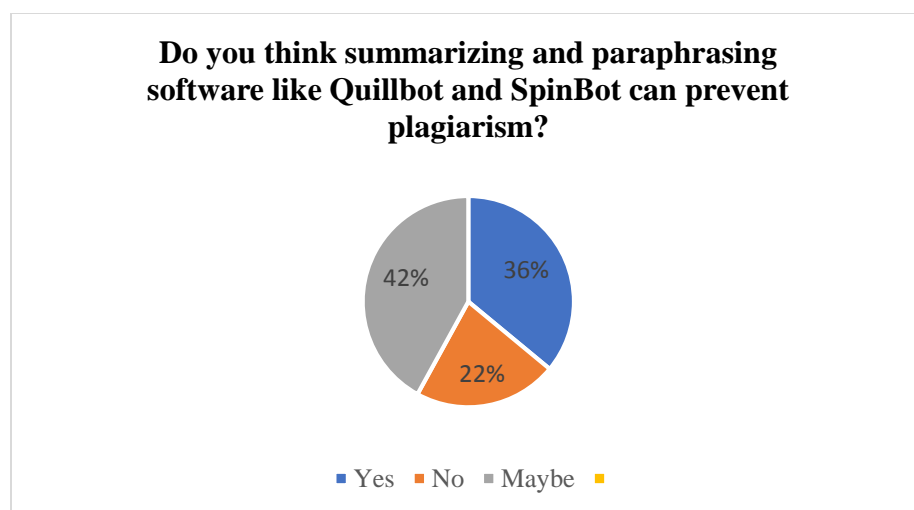
22%

■ Yes ■ No ■ Maybe ■

Figure 7. The Benefit of AI Support Preventing Plagiarism (N=32)

7.   Discussion

This paper aimed to assess the use of AI-based summarizing and paraphrasing tools for a CLIL intercultural communication academic writing class based on the perceptions of the students who used them. Results showed that the Quillbot and SpinBot tools had both advantages and limitations. Quilbot was found to be the better of the two software applications and produced better grammar scores when compared to summaries written with no AI support. SpinBot showed no statistical differences and its help in improving a student's grammar was limited. In terms of the style of the student's summaries (readability, structure) both of the online tools provided no real impact in improving students' summary scores. SpinBot was found to reduce scores in some cases.

Results were also inconclusive as to whether the AI tools can help students avoid plagiarizing the content they are summarizing and paraphrasing. No statistical differences were found in the Plagscam data that scanned the summaries for matching content with the summaries that were written with and without the support of the AI tools.

Students found the tools to be helpful, especially the Quillbot application, and suggested that the tools can be potentially helpful in their English academic writing studies. Paraphrasing was found to be the most difficult aspect of writing their summaries without the use of AI support but this was not a concern when they used these tools. This would suggest that the tools fulfilled their job of helping the students paraphrase their article content. However, many students expressed their concerns about the accuracy of the results that these tools produced and had to edit and change the content as it produced grammatically inaccurate and unnatural sentences. The formality of the sentences produced was also questioned by students which would suggest that context and pragmatic considerations can be raised when adopting these tools within the classroom.

Although the findings within this case study were mixed, the successful implementation of new instructional technologies is closely related to the attitudes of the teachers who lead the lesson and the instructions given on how to use the tools. Based on the findings in this study the tools can be a useful aid for teachers and students to accomplish specific writing tasks such as summarizing and paraphrasing which are essential skills in academic writing. However, they must not be a replacement for good teaching but as a teacher and student companion.
The findings within this paper have pedagogical implications on how to adopt AI-based tools in university EFL academic writing classes.

As many universities around the world have adopted to use of AI technology within their English language learning classrooms it is hoped the findings in this study can be beneficial to teachers who are looking for ideas on how to integrate such tools and what expectations and results from the tools can provide them. The results of this paper were based on the findings and work done within AI tools and apps. The accuracy and reliability of these apps and the data that they produced can be questioned; however, this paper intended to assess the benefits of using AI-based tools within the English language learning classroom and the possible implications of these results. When asking the question, 'Can teachers rely on and use these tools as a replacement for teaching, evaluating, and checking for plagiarism? then the answer is 'no.' These tools can, however, be used as an aid for teachers (although a potentially expensive one) and the summarizing and paraphrasing tools evaluated in this paper can help students understand the principles of summarizing and paraphrasing and help improve their vocabulary through the number of synonyms offered in the results.

8.   Conclusion

This study aimed to assess the benefits of using AI tools for summarizing and paraphrasing class assignments in a CLIL Intercultural Communication Academic Writing course. Findings showed that the use of Quillbot improved students' grammar, but this was not the case with SpinBot. Both AI-based tools did not improve the writing style scores of the students when compared to their summaries without AI support. In addition, results were inconclusive on whether or not these tools can prevent plagiarism which paralleled student assertions within the after-summary assignment survey.

The use or dependence on such tools to produce academic writing content, therefore, needs to be treated with caution. Online paraphrasing tools can provide the potential for students to submit work they have not directly written themselves but done 'artificially.' The quality of the output generated can be questioned and the free versions of this kind of software can create problems rather than solve them. As the students suggested, some of the changes the software makes may be inaccurate, and change the nuance or meaning of what the student wishes to convey. Depending on a student's grammar and vocabulary knowledge, the ability to spot mistakes (unnatural sentences) and edit their work will differ. This kind of software, therefore, should be used as an aid rather than a replacement of the writer in compiling essays.

This study, however, only examined AI-based summarizing and paraphrasing tools in addition to apps that can assess a student's writing and evaluate the writing for plagiarism. There are a huge number of AI tools ranging in quality, price, and technological advancement which can be used personally or adopted by educational institutions and integrated into their curriculums. COVID-19 has created the need for AI and if utilized properly, will bring huge advantages within the education sector through online classes which can be efficient and effective. In the future, the new normal may result in a hybrid learning environment where humans and robots can work together in ways never imaginable.

References

Bailey, Daniel; Southam, Ashleigh, & Costley, Jamie. (2021). Digital storytelling with chatbots: mapping L2 participation and perception patterns. *Interactive Technology and Smart Education, 18*(1), 85-103.

Bailin, Alan. (1987). Artificial intelligence and computer-assisted language instruction: A perspective. *CALICO Journal, 5*(3), 25-45.

Basilaia, Giorgi, & Kvavadze, David. (2020). Transition to online education in schools during a SARS-CoV-2 Coronavirus (COVID-19) pandemic in Georgia. *Pedagogical Research, 5*(4), em0060.

Burleson, Winslow, & Lewis, Armanda. (2016). Optimists' creed: Brave new cyberlearning, evolving utopias (Circa 2041). *International Journal of Artificial Intelligence in Education, 26*(2),796-80

Coyle, Do.; Hood, Philip, & Marsh, David. (2010). *Content and language integrated learning*. Cambridge: CUP.

Educause. (2018). Horizon report: 2018 higher education edition. Retrieved from EDUCAUSE Learning Initiative and The New Media Consortium website: https://library.educause.edu/~/media/files/library/2018/8/2018horizonreport.pdf

El Shazly, Reham. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems, 38*(3), e12667.

Gallacher, Andrew; Thompson, Andrew, & Howarth, Mark. (2018). "My robot is an idiot!" – Students' perceptions of AI in the L2 classroom. In P. Taalas, J. Jalkanen, L. Bradley & S. Thouësny (Eds), *Future-proof CALL: Language learning as exploration and encounters – short papers from EUROCALL 2018* (pp. 70-76).

Istenic, Andreja; Bratko, Ivan, & Rosanda, Violeta. (2021). Pre-service teachers' concerns about social robots in the classroom: A model for development. *Education and Self Development, 16*, 60-87.

Kaplan, Andreas, & Haenlein, Michael. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*, 15-25.

Kaban, Ayseguk Liman, & Ergul, Isil Boy. (2020). Teachers' attitudes towards the use of tablets in six EFL classrooms. In P. Eva (ed.), *Examining the Roles of Teachers and Students in Mastering New Technologies* (pp. 284-298). Pennsylvania: IGI Global.

Keck, Casey. (2006). The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing 15*(4), 261-278.

Keck, Casey. (2014). Copying, paraphrasing, and academic writing development: A re-examination of L1 and L2 summarization practices. *Journal of Second Language Writing*, *25*, 4-22.

Matthews, Clive. (1993). Grammar frameworks in intelligent call. *CALICO Journal*, *11*(1), 5-27.

Mulenga, Eddie, & Marbán, Jose. (2020). Is COVID-19 the gateway for digital learning in mathematics education? *Contemporary Educational Technology, 12*(2), ep269, 1-11.

Naciri, Aziz; Baba, Mohamed; Baba, Amine; Achbani, Achbani, & Kharbach, Ahmed. (2020). Mobile learning in higher education: Unavoidable alternative during COVID-19. *Aquademia, 4*(1), ep20016.

Nilsson, Nils. J. (2011). *The quest for artificial intelligence: A history of ideas and achievements.* Cambridge University Press.

Niño, Ana. (2009). Machine translation in foreign language learning: Language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL, 21*(02), 241-258.

Ohmori, Ai. (2014). Exploring the potential of CLIL in English Language Teaching in Japanese universities: An innovation for the development of effective teaching and global awareness. *The Journal of Rikkyo University Language Center Bulletin*, *32*, 39–51.

Pathways to Academic English. (2021). The Center for Culture and Language Education English Section, Institute for Excellence in Higher Education, Tohoku University, Tohoku University Press.

Prensky, Mark. (2007). How to teach with technology: Keeping both teachers and students comfortable in an era of exponential change. *Emerging Technologies for Learning*, *2*, 40-46.

Prensky, Mark. (2008). Backup education? Too many teachers see education as preparing kids for the past, not the future. *Education Technology, 48*, 1-3.

Rogerson, Ann M., & McCarthy, Grace. (2017). Using internet based paraphrasing tools: Original work, patchwriting or facilitated plagiarism? *International Journal for Educational Integrity*, *13*(2), 1-15.

Salaberry, M. Rafael. (1996). A theoretical foundation for the development of pedagogical tasks in computer mediated communication. *CALICO Journal, 14*(1), 5-34.

Shi, Ling. (2012). Rewriting and paraphrasing source texts in second language writing. *Journal of Second Language Writing*, *21*,134-148.

Sintema, Edgar John. (2020). E-learning and smart revision portal for Zambian primary and secondary school learners: A digitalized virtual classroom in the COVID-19 era and beyond. *Aquademia, 4*(2), ep20017.

Tafazoli, Dara; María, Elena Gomez, & Huetas Abrill, Cristina Aranzaza. (2019). Intelligent language tutoring system: Integrating intelligent computer-assisted language learning into language education. *International Journal of Information and Communication Technology Education, 15*(3), 60-74.

Wartman, Steven. A., & Combs, Donald. (2018). Medical education must move from the information age to the age of artificial intelligence. *Academic Medicine*, *93*(8), 1107-1109.

Yin, Jiaqi; Goh, Tiong-Thye; Yang, Bing, & Yang, Xiaobin. (2021). Conversation technology with micro-learning: The impact of chatbot-based learning on students' learning motivation and performance. *Journal of Educational Computing Research, 59*(1), 154-177.

Author Note

Barry Kavanagh is an associate professor at Tohoku University, Japan. His research interests include CLIL, bilingualism, second language acquisition and computer mediated communication and he has published widely in these areas. His publications include:"A study on how cultural and gender parameters affect emoticon distribution, usage and frequency in American and Japanese online discourse" (2021) in *Approaches to Internet Pragmatics: Theory and Practice* (edited by Chaoqun Xie, Francisco Yus and Hartmut Haberland) and "Developing intercultural awareness and language learning through a CLIL course on karate and Japanese culture" (2023) in *Re-envisioning Language Teaching & Learning in Asia* (edited by Theron Muller, John L. Adamson, Steven Herder, and Philip Shigeo Brown).

He is the Vice President of the J-CLIL pedagogy association and chair of the J-CLIL Tohoku chapter.