EXPLORING THE IMPACT OF AI-POWERED COLLABORATIVE AND INTERACTIVE NLP APPS ON EFL TEACHING IN THE POST-COVID-19 ERA

Coffi Martinien ZOUNHIN TOBOULA

English Department, University of Abomey Calavi (UAC), Benin
Laboratoire du Groupe de Recherche sur l'Afrique et la Diaspora (GRAD)

ABSTRACT

The COVID-19 pandemic has disrupted traditional language learning, leading to a shift towards online teaching and requiring new approaches to language education. This study examines the effectiveness of AI-powered collaborative and interactive Natural Language Processing (NLP) applications on English as a Foreign Language (EFL) instruction in a post-COVID-19 online education environment. The study used a mixed-methods approach, incorporating statistical and in-depth qualitative data gathering and processing strategies. EFL teachers and students from the University of Abomey-Calavi (UAC) in Benin were surveyed, interviewed, and observed during online language learning sessions. The data were analysed using both descriptive and inferential statistics. The study employed questionnaire surveys to analyse quantitative data and used the thematic (content) analysis method to isolate the most important trends and themes hidden within the qualitative data collected through semi-structured interviews and online class observations. Results showed the challenges and opportunities of using AI-powered collaborative and interactive language learning in EFL teaching, the learning methodologies and assessment approaches used in AI-enabled collaborative e-learning, the role of technology in supporting pervasive learning, and the impact of professional development for teachers in ICT on integrating AI-assisted collaborative e-learning in EFL instruction. The findings offer new perspectives on the effects of AI-supported collaborative and interactive language learning on EFL teaching and its implications for EFL teachers and students in the post-pandemic era.

KEYWORDS

AI-powered, Collaborative and Interactive Language Learning, NLP Apps, Post-COVID-19, Online EFL Teaching

1. INTRODUCTION

Teaching English as a Foreign Language (EFL) has become increasingly challenging due to the unparalleled impediments presented by the COVID-19 pandemic. The traditional classroom-based approach to language learning has been disrupted, leading to a rapid shift toward online learning. In this context, AI-powered collaborative and interactive language learning is becoming increasingly relevant to enhance EFL teaching in the post-COVID-19 era (Abd-Alrazaq et al.,2020; Warschauer, 1996 and 1997).

Artificial Intelligence in language learning has recently gained traction (Holmes, Persson, Chounta, Wasson, and Dimitrova, 2022). AI-powered language learning systems are designed to provide personalised learning experiences for students, enabling them to learn at their own pace

David C. Wyld et al. (Eds): DBML, CITE, IOTBC, EDUPAN, NLPAI -2023
pp.171-193, 2023. IJCI – 2023 DOI:10.5121/ijci.2023.120213

and in their way. AI-powered language learning systems can also facilitate collaborative and interactive learning, allowing students to work to find solutions to issues and have meaningful interactions (McArthur, Lewis, and Bishary, 2005; Ouyang and Jiao, 2021).

Theoretical frameworks such as Connectivism (Siemens, 2005), social constructivism (Vygotsky and Cole, 1978), and cognitive load theory (Sweller, 1988) provide a strong foundation for this study. These theories posit that language learning is a social and cultural process in which learners engage with one another to negotiate meaning and build their understanding of the target language. AI-powered collaborative and interactive language learning tools, such as chatbots, virtual assistants, and gamified learning environments, provide opportunities for students to engage in authentic, task-based interactions and to receive real-time feedback from their peers and teachers. Connectivism emphasises the importance of networks and connections in learning, while social constructivism highlights the role of social interactions in the learning process (Siemens, 2005; Vygotsky and Cole, 1978). Cognitive load theory explains how learners process information and suggests that learners learn best when instructional materials are well-designed (Sweller, 1988, Sweller, Van Merrienboer, and Paas, 1998). Furthermore, these theories can inform the design of AI-powered language learning systems, allowing for more effective and efficient learning.

The COVID-19 pandemic has posed new challenges for traditional language education, and the impact of AI-powered collaborative and interactive language learning on EFL teaching is yet to be fully understood. There is a need to examine the effectiveness of this teaching approach and its impact on EFL instruction in the post-pandemic online teaching environment.

It is hypothesised that AI-powered collaborative and interactive language learning can enhance EFL teaching in the post-pandemic online environment (Chen, Chen, and Lin, 2020). Using AI-enabled collaborative e-learning can lead to increased engagement in EFL teaching, improved learning outcomes for EFL students, and increased teacher satisfaction. Moreover, professional development for teachers in ICT (Hennessy et al.,2021) can be beneficial for integrating AI-assisted collaborative e-learning in EFL instruction and for empowering educators by equipping them with the relevant skills to use AI-enabled tools in their EFL teaching effectively.

Based on the objectives and hypotheses of the present study, the following research questions have been developed to provide a framework for this research investigation.

1. How does AI-powered collaborative and interactive language learning impact EFL teaching and learning in the post-COVID-19 era, including challenges, opportunities, and implications?
2. How do AI-enabled collaborative e-learning methods, assessment approaches, and technology support impact EFL teaching in the post-COVID-19 era?
3. What is the effect of professional development for teachers in ICT on integrating AI-assisted collaborative e-learning in EFL instruction, and what is its impact on EFL teaching and learning in the post-pandemic era?

A research approach that combined quantitative and qualitative data collection and analysis methods was employed in this study to produce a comprehensive analysis. Surveys, semi-structured interviews, and observation of online language learning sessions were conducted with EFL teachers and students recruited from the University of Abomey-Calavi (UAC) in Benin. Descriptive and inferential statistics were used to analyse the quantitative data. Conversely, the qualitative data obtained from online class observations and semi-structured interviews were subjected to thematic (content) analysis to reveal prominent themes and patterns. The findings of this study provide insights into the impact of AI-supported collaborative and interactive language

learning on EFL teaching and its implications for EFL teachers and students in the post-pandemic era.

2. REVIEW OF RELEVANT WORK

For many years, the influence of Artificial Intelligence (AI) on language learning has been a subject of fascination. In recent times, AI-driven collaborative and interactive language learning has been gaining momentum, and its capacity to improve English as a Foreign Language (EFL) teaching in the post-COVID-19 era is an area of research that deserves attention. This section provides an overview of the theoretical framework of the study. It examines relevant studies, emerging research questions, and key points from sources that have tackled similar issues.

2.1. Theoretical Framework

Artificial intelligence (AI) has become increasingly prevalent in language learning in recent years, particularly in response to the COVID-19 pandemic. AI-powered collaborative and interactive language learning tools can create a learning environment that is more engaging and tailored to the learners' interests, which can improve their language proficiency outcomes. This work explores the theoretical framework that underpins the impact of AI-powered collaborative and interactive language learning on enhancing English as a Foreign Language (EFL) teaching in the post-COVID-19 era.

Connectivism is a philosophy of learning that recognises the power of networks and connections in creating and sharing knowledge (Siemens, 2005), as exemplified by how people can collaborate and learn together through online platforms and social media. According to this theory, learning is an ongoing process of connecting and constructing knowledge through interactions with others. Connectivism suggests that learners develop their personal comprehension of the target language by actively exploring and creating connections between different concepts and notions. Goldie (2016) examines the concept of connectivism and its potential application in medical education. The author discusses this concept as a network learning theory developed for e-learning environments. The principles of connectivism, as presented by Goldie (2016), include the idea that learning and knowledge come from a diversity of opinions and that learning involves connecting specialised nodes and information sources. The author concludes that connectivism provides a valuable lens for understanding and managing teaching and learning with digital technologies, but further development and testing are necessary. Therefore, AI-powered language learning tools can be designed to facilitate this process by providing learners with opportunities to engage in authentic, task-based interactions with their peers and teachers.

Another critical theory that informs our understanding of the impact of AI-powered collaborative and interactive language learning is social constructivism. It is an educational theory that places emphasis on the role of social interactions in learning (Vygotsky and Cole, 1978). Social constructivism posits that learning is a social process that occurs through interaction and collaboration with others. The theory of social constructivism proposes that acquiring knowledge is a collaborative and social endeavour involving engaging and interacting with others. In the context of language learning, social constructivism suggests that students will benefit from engaging in collaborative and interactive language learning activities, such as those facilitated by AI-powered language learning tools. For example, a study by Lee et al. (2022) found that students who participated in collaborative language learning activities using AI-powered tools showed more remarkable language proficiency improvement than those who participated in individual language learning activities. Chatbots can be employed to give learners instantaneous

feedback on their language use, while virtual assistants can be used to provide learners with tailored language learning experiences.

Kalina and Powell (2009, p.243) suggest that successful teaching and learning in the classroom are reliant on the use of constructivist approaches and tools. They identify two main types of constructivism: Cognitive or individual constructivism, grounded on Piaget's (1953) notion that knowledge is formed through a self-driven process, and social constructivism, which is based on Vygotsky's (1962) theory of knowledge being built through interaction with the teacher and other students. To be effective, Kalina and Powell (2009, p.243) argue that teachers must be familiar with both theories and apply constructivist teaching methods, strategies, tools, and practices. They back up their claim by contrasting Piaget's (1953) theory, which focuses on the individual and how they construct knowledge, with Vygotsky's (1962) theory, which emphasises the role of language in the development and the interaction between the individual and the social environment.

Cognitive Load Theory (CLT) is a learning concept that focuses on the limitations of human working memory and how it affects learning new information. It provides insight into how learners process information and suggests that instructional materials should be designed to optimise learning (Sweller, 1988; Sweller, Van Merrienboer, & Paas, 1998). This theory suggests that learners should be given instructional materials that are tailored to their specific needs and capabilities. According to Sweller, Ayres, and Kalyuga (2011), CLT divides knowledge into biologically primary and secondary knowledge, with the latter being the main focus of educational institutions. The theory emphasises the importance of understanding human cognitive architecture for instructional design and classifies cognitive load into intrinsic and extraneous load based on their impact on working memory. Intrinsic load is composed of elements essential to learning, while extraneous load is composed of features that are not necessary for learning and are a function of instructional procedures.

Another important theoretical framework is the self-determination theory (SDT), which posits that motivation is considered a crucial factor in the learning process, with evidence suggesting that students are more inclined to engage, be motivated, and participate when they have autonomy, competence, and relatedness in their learning experiences. In the context of AI-powered language learning, SDT advocates for collaborative work among students, as it allows for a sense of shared purpose and fosters an environment of collective learning and interaction with AI-powered language learning tools in a way that supports learners' autonomy and competence.

To sum up, a theoretical framework based on social constructivism and self-determination theory can be used to understand the effects of AI-powered collaborative and interactive language learning on EFL teaching in the post-COVID-19 era. These theories emphasise the significance of collaboration and interaction in language learning and motivation's role in improving language proficiency results. Further research is necessary to investigate how AI-powered collaborative and interactive language learning tools can support these essential elements and improve EFL teaching in the post-COVID-19 era.

2.2. Consideration of other Related Work

The COVID-19 pandemic has spurred Artificial Intelligence (AI) promotion in Benin Republic. On January 5, 2021, a conference on AI was held at the Benin Excellence Library in Godomey, in the Atlantique region of Benin. It was led by Herbert Brian Whannou, a statistician who graduated from the University of Paris Diderot and the Polytechnic Institute of Paris. The conference's main topic was "Creating a Machine Learning Model: Best Practices and Common

Mistakes to Avoid." This situation sparked interest, leading to several start-ups creating AI-based programs and the first African edition of the Summer School on Artificial Intelligence (EEIA 2021) held in Godomey from July 19 to August 13, 2021. The first edition's success prompted the second edition (EEIA 2022) to be held from July 18 to August 12, 2022. The Vallet Foundation organised it in partnership with the NGO Benin Excellence and the United Nations Development Programme (UNDP). The event was the most prominent educational AI gathering in French-speaking Africa, attracting around 5,000 visitors from French-speaking African countries, with ten students being awarded scholarships. During the EEIA 2022, 100 young people received practical training in AI basics such as programming, machine learning, and embedded electronics. These courses lasted four weeks and were free, with limited spots available through a three-stage selection process based on an academic level, personal project, and motivation. This situation highlights the significance of AI in Beninese education. It explains why a national AI and mega-data strategy was established in Benin at the Council of Ministers meeting on January 18, 2023 (MND, 2023).

This national plan for artificial intelligence and mega data has been established to promote the use of digital technology for economic and social progress. The strategy involves stakeholders and adopts a holistic approach to address the country's needs, particularly in education, health, agriculture, living environment, and tourism. It is divided into four programs, implemented in three phases over five years, with 123 actions aimed at impacting the public and private sectors. The goal is to make Benin a leader in using AI and big data by 2027 and attract investment and talent to the country. As a result, the education sector will become a hub for AI-based training programs and innovation.

As Whannou (2021) pointed out, machine learning is a branch of data science that enables systems to enhance performance through the experience without requiring specific programming. This field has opened doors to collaborative and interactive learning, considered among the most effective methods to involve and captivate EFL learners. With the advent of COVID-19, various AI-based training programs have been promoted in Benin, including *Google Translate*, *Lingvist*, *Duolingo*, Siri, Cortana, Rosetta Stone, Mondly, and *Babbel*, among others. All of them are based on NLP (Natural Language Processing), a technology that harnesses the power of Artificial Intelligence.

Implementing these programs may improve the likelihood of success in foreign language learning as they rely on an individual's cognitive and memory abilities, intelligence, and capacity to store information. AI is no longer just science fiction; the future is now! The future of AI is now, and its impact on education and other fields is rapidly growing.

As it is crucial to familiarise yourself with some foundational concepts in the field to gain a deeper understanding of Artificial Intelligence, we will draw insights from Schmaus (2022), an expert at Talkwalker, a company specialising in developing AI-based programs. This expert provided a quick overview of crucial AI terms to enhance knowledge in this area.

According to Schmaus (2022), *Artificial Intelligence (AI)* was first introduced at the Dartmouth Conference in 1956. It refers to computer systems that mimic human learning and problem-solving functions. Schmaus (2022) defined *Artificial Intelligence* as Machines designed to learn, solve problems, and perform tasks using human mental processes as models. He argues that AI automates complex and repetitive tasks while freeing humans to focus on more abstract tasks beyond a machine's capabilities. Dobrev (2012) defines *Artificial Intelligence (AI)* as a technology that separates knowledge from intelligence, a program that achieves a level of performance on par with that of human capabilities in any environment. This definition is based on three assumptions: every calculation device can be modelled by a program, AI is a step device

that inputs and outputs the information, and AI is in an environment that provides information and is influenced by its output.

Artificial Intelligence (AI) includes a range of techniques and approaches, including sentiment analysis, predictive analytics, machine learning, reinforcement learning, deep learning, and supervised/unsupervised learning. *Sentiment analysis* combines natural language processing, computational linguistics, and textual analysis to identify and extract subjective information from content. *Predictive analytics* leverages previous data to forecast future trends or outcomes through *machine learning*, statistics, and data mining. AI improves itself through experience or learning, and *deep learning* is the most advanced form. *Supervised and unsupervised learning* are two methods of educating AI, with the former using human-labeled datasets and the latter allowing the AI to assign categories to the results (Schmaus, 2022). Additional terms similar to those mentioned below have also been given definitions by Schmaus (2022).

- *Digital Assistants*: Smartphone software (such as Siri, Google Assistant, Cortana, and Alexa) designed to respond to voice requests and allow for hands-free use of phone functions.
- *Big Data*: A term describing the exponential growth of data, requiring computing power beyond current software capabilities and specific requirements for data flow, collection, storage, and analysis.
- *Chatbot*: Artificial intelligence programs mimicking human conversation, used in various applications such as customer service, instant messaging, and virtual assistants.
- *ChatGPT*: OpenAI's breakthrough natural language conversational tool capable of generating text using advanced AI, introduced in November 2022.
- *Human-Computer Interaction*: A discipline studying the interaction between humans and computer technology, combining design, psychology, and computer science
- *Collaborative Apps* are designed to facilitate communication and collaboration between users while *Interactive Apps* are designed to provide an engaging and immersive user experience and the two concepts are not necessarily interchangeable.
- *Algorithm*: An algorithm is a collection of predetermined protocols for carrying out a sequence of simple to increasingly complex actions, including calculations, data processing, or the automation of repetitive tasks.

Speech recognition allows a machine to interpret human speech and convert it into a computer-readable format. – *Artificial neural networks* are modelled after the human brain and are designed to create more efficient machine learning systems. *A robot* is a device that performs repetitive tasks automatically. – *Computer learning theory* studies the design and analysis of machine learning algorithms.

*Automatic **N**atural **L**anguage **P**rocessing* (A**NLP**) utilises machine learning to equip computers with the ability to comprehend natural language in written or spoken form. *Large Language Models* (LLM) are advanced linguistic models developed to anticipate the next word in a sentence by processing extensive linguistic patterns and performing tasks such as translation, summarisation, and answering questions.

Jiang (2022) explores how AI technology is used to enhance EFL teaching and learning. The author provides an overview of six main forms of learner-facing and teacher-facing (Baker and Smith, 2019, as observed in Jiang, 2022, p.4). They include AI applications in the EFL context, including Automatic Evaluation Systems (AESs), Neural Machine Translation (NMT) tools, Intelligent Tutoring Systems (ITSs), AI Chatting Robots (AI chatbots), Intelligent Virtual Environment (IVE), and affective computing in ITSs. The study highlights the current lack of research on the application of affective computing in the EFL context and the need for further

exploration of AI's pedagogical and ethical implications in the EFL context. Jiang (2022) concludes by discussing challenges from both technical and teacher perspectives and future research directions.

Though Jiang's (2022) study mentioned the potential benefits of AI in EFL in terms of promoting adaptive learning and automating tasks such as assessment and feedback, it did not mention any potential drawbacks or limitations. For example, while AI-powered systems can automate tasks such as assessment and feedback, they may also limit the opportunity for human interaction and critical thinking skills. Additionally, while deep learning-based AI techniques have shown potential in the EFL context, they also come with challenges, such as analysing multi-modal signals and considering student emotions and moods. Another essential issue is the ethical implications and risks of using AI in EFL education. However, it is good to note that Jiang's (2022) study mentioned a lack of research from a teacher's perspective to fully understand the impact of AI on EFL and address any potential risks.

Junaidi's (2020) work reports on a study that aimed to measure the effectiveness of artificial intelligence (AI) in developing the speaking abilities of EFL students in classrooms. The study used the Lyra Virtual Assistant (LVA) app, which was selected due to its unique features, ease of use, and low cost. The sample consisted of 65 students from two seventh-grade classes who were divided into an experimental and a control group. A quasi-experimental approach was taken, incorporating pre-and post-tests to measure the students' speaking performance in pronunciation, grammar, vocabulary, and fluency. The results showed that the experimental group, which used LVA, significantly improved their speaking performance compared to the control group, which used conventional methods. The authors conclude that LVA is an AI-powered application that is useful for EFL students to enhance their speaking abilities.

A critical evaluation of this study highlights several limitations that need to be considered when interpreting its results. Firstly, the sample size of 65 students is relatively small and may not be representative of larger populations, thereby reducing the generalizability of the results. Secondly, the use of a quasi-experimental design raises questions about causality and the potential for omitted variable bias. Using a true experimental design with the random allocation of participants to either the experimental or control group could have improved the study's robustness. Furthermore, the study's attention to only four components of speaking skills may not fully encompass the nuances of spoken language proficiency. A more comprehensive assessment that includes a broader range of speaking skills would provide a more nuanced understanding of the effectiveness of the Language Visualization Algorithm (LVA). Finally, the study was conducted in a single secondary school in Indonesia, and its external validity, or generalizability to other EFL contexts, is unclear. The study results could be expanded to various cultural and educational environments for further investigation. Additionally, the control group in the study used conventional teaching methods. Still, the specifics of these methods are not described, making it difficult to compare the results between the experimental and control groups.

Overall, Junaidi's (2020) study provides some initial evidence of the potential of AI apps like LVA to improve students' speaking performance in an EFL context. However, the study's limitations suggest further research to validate the results and examine the long-term effectiveness of AI in language learning.

Pokrivčáková's (2019) research focuses on integrating artificial intelligence (AI) in foreign language education. It begins by discussing the impact of AI in various areas of our daily lives. It provides a background on AI definition and its evolution after highlighting the challenges in defining this theory and the different perspectives in its definition. Subsequently, it outlines the modifications that resulted from the discussed changes by applying AI-powered tools,

specifically Intelligent Computer-Assisted Language Learning (ICALL) as a subset of computer-assisted language learning (CALL). The author summarises existing research on AI-powered tools for foreign language education ["machine learning, adaptive learning, natural language processing, data mining, crowdsourcing, neural networks, and algorithms" (Pokrivčáková's, 2019, p.135)] and their results. She argues for the need to prepare foreign language teachers effectively to integrate AI into their teaching.

Pokrivčáková's (2019) research provides an overview of AI in foreign language education and its potential impact. However, it only summarizes existing research on the application of AI-powered tools and does not give a comprehensive overview of the field. The results of the existing research mentioned in the study are also scarce, which limits the validity of the conclusions drawn. However, the author acknowledged that and called for further investigation and exploration. Moreover, the study only considers AI technologies. It does not examine other technology forms that could benefit foreign language education, such as gamification or virtual reality. It does not consider other factors that could impact the preparation of foreign language teachers. In short, the study conclusions and recommendations would have benefited from a more thorough examination of the field and a broader consideration of technology in foreign language education.

Abalkheel's (2022) study explores the challenges of online EFL learning in Saudi Arabia amid the post-COVID-19 era. The author argues that combining Bloom's Taxonomy and Artificial Intelligence (AI) can provide strategies to overcome these challenges and provide effective teaching outcomes. The study discussed how AI could be used to create automated formative assessments for EFL learners, provide personalised feedback and generate customised learning experiences. Furthermore, the author offers a framework for incorporating AI into the existing EFL pedagogical structure, which will enable greater instructional effectiveness.

This study is limited in scope, as it only focuses on how AI and Bloom's Taxonomy can be used to improve the teaching and learning of EFL in Saudi Arabia. The study does not consider how other countries might benefit from this approach, nor does it discuss the potential risks of introducing AI into EFL teaching. Apart from supporting evidence, the author fails to demonstrate the effectiveness of these strategies, warranting the need for further research to assess their full efficacy.

In light of the aforesaid, this investigation appears rational by proposing to corroborate the several postulations raised by each of these studies while simultaneously striving to fill in any gaps left by them. By instituting an investigative methodology attuned to this outlook, this study has endeavoured to respond to its various basic questions to validate the postulated hypotheses congruent with its objectives.

2.3. Methodology

The present study employed a mixed-methods design that integrated quantitative and qualitative data collection and analysis techniques. The research design was chosen to provide a more comprehensive and in-depth understanding of the impact of AI-powered collaborative and interactive language learning on EFL teaching in the post-COVID-19 era.

The study received ethical approval from the University of Abomey-Calavi's Institutional Review Board (IRB) to ensure that the study was conducted in an ethical and responsible manner, with the privacy and rights of participants protected. The research took place in the context of the English Department and the Beninese Center for Foreign Languages (CE.BE.LA.E) of the

University of Abomey-Calavi (UAC) in Benin, where EFL teachers and students were recruited to participate in the study.

The study participants consisted of 30 EFL teachers and 431 students from the English Department (18 EFL Teachers and 327 EFL learners) and the Beninese Center for Foreign Languages (12 EFL Teachers and 104 EFL learners) of the University of Abomey-Calavi. A purposive sampling method was used to recruit the participants, and all participants consented to participate in the study.

Table 1: Demographic Distribution of Participants by Gender, Institution, and EFL Proficiency in AI Tools.

	Gender	**EFL Teachers**		**EFL Learners**	
		P	**N/P**	**P**	**N/P**
English Department (EngD)-UAC					
	Male	4	7	-	197
	Female	1	6	-	87
	ND	-	-	-	43
Subtotal		**18**		**327**	
Beninese Center for Foreign Languages (CE.BE.LA.E)-UAC					
	Male	2	5	-	69
	Female	3	2	-	27
	ND	-	-	-	8
Subtotal		**12**		**104**	
Note: ND= Non-disclosed, P= Proficient in EFL AI-powered tools, N/P= Non-Proficient in EFL AI-powered tools					

A hundred AI-based programs with gamification and feedback systems were selected as part of the study preparation to create collaborative and interactive learning experiences in English as a Foreign Language (EFL) instruction. With the assistance of 10 of the 30 expert teachers involved in the project that had experience in teaching EFL through AI-powered tools, the two best programs were chosen from the following five categories: AI-Powered Virtual Classrooms, AI-Powered Chatbots, AI-Powered Sentiment Analysis Tools, AI-Powered Natural Language Processing (NLP) Tools, and Deep Learning Tools. After a thorough evaluation, the three most suitable categories for EFL teaching and practice, AI-Powered Virtual Classrooms, AI-Powered Natural Language Processing (NLP) Tools, and AI-Powered Sentiment Analysis Tools, were selected. Programs such as Classcraft, Google Classroom (AI-Powered Virtual Classrooms), Lexalytics and VADER (AI-Powered Sentiment Analysis Tools), The Natural Language API offered by Google Cloud and the Natural Language Understanding platform provided by IBM Watson [AI-Powered Natural Language Processing (NLP) Tools] were approved by a committee of teachers in charge of validating EFL instructional materials for use in sub-Saharan Africa, specifically Benin. This committee was responsible for evaluating and endorsing the research tools used in the study. The committee chose Classcraft and Google Classroom (AI Virtual Classrooms), Google Cloud Natural Language API, and IBM Watson Natural Language Understanding (AI-Powered NLP Tools) as the best AI technologies for evaluating and implementing EFL learning in sub-Saharan Africa, specifically in Benin. The study ultimately utilised only ***Google Classroom*** and ***IBM Watson Natural Language Understanding*** tools due to budget constraints as they offered free plans or a free trial. Google Classroom is a free platform for educational institutions, while IBM Watson Natural Language Understanding provides a free plan along with paid options. On the other hand, Classcraft is a paid program but offers a limited-time trial.

Two separate groups were employed for the aims of this research. One group, the experimental group (ExpG), consisted of 327 EFL learners from the English Department and 104 from CE.BE.LA.E and was exposed to AI-powered Collaborative and Interactive Language Learning using Google Classroom and IBM Watson Natural Language Understanding tools. The control group (ConG), composed of the remaining classmates (203 from the English Department and 108 from CE.BE.LA.E), did not use any AI-powered technology. The study participants from the English Department and CE.BE.LA.E were first-year students with mostly a post-beginner level (CEFR A1-A2+), as determined by the pre-test results. Due to limited classroom space, the classes were divided into two groups, and the rolling class technique was used. The study involved particular EFL face-to-face and distance learning classes using AI technology for 3 out of 5 days and traditional teaching methods for the other days with peers from the control groups. Both groups were given a pre-test and post-test assessments based on the evaluation criteria for the Cambridge English exams that align with the Common European Framework of Reference (CEFR) for Languages (Cf. Table 2). Learners were motivated to attend private classes with their own smartphones and computers or those available at the Beninese Center for Foreign Languages (CE.BE.LA.E). Free internet access was provided to all participants to enhance motivation.

Table 2. Criteria for Cambridge English exams, based on the CEFR

Skills	Evaluation Criteria
Grammar and Vocabulary	The aptitude to employ a variety of grammatical constructions and lexicon precisely and fittingly. Knowledge of common idiomatic expressions.
Reading and Writing	Ability to understand and produce written texts in various genres and styles. Use of a range of reading and writing strategies to comprehend and produce written texts.
Listening	The capacity to comprehend spoken English in diverse circumstances and settings and the skill to recognize principal concepts and precise particulars in oral materials. Ability to follow the development of arguments and narratives.
Speaking	Ability to communicate effectively in spoken English. Use of appropriate pronunciation, grammar, and vocabulary. Use of appropriate discourse markers and connectors to link ideas and maintain coherence in spoken discourse.

The 12-week EFL experiment involved using AI technology for 45-minute face-to-face and 1.5-hour online classes per week. Classes took place at the English department and language laboratory of the Beninese Center for Foreign Languages. The experimental group participated in the special AI-based courses while the control group continued with their regular classes. Both groups took a pre-test and post-test using the Cambridge English test to measure language proficiency, including a speaking test that involved talking about one of their typical days and participating in an interactive group discussion. The assessment standards were aligned with the CEFR criteria and adapted from the model developed by Cambridge. Only the experimental group received AI-based training materials selected by 30 teachers, while the control group continued with regular grammar, introductory linguistics, civilisation, and literature (African, British, and American) classes. This research was carried out from September 6, 2021, to January 28, 2022, in the 2021/2022 academic year.

Data collection was performed using surveys, semi-structured interviews, and observation of online language learning sessions. The surveys were administered to both EFL teachers and students, and the semi-structured interviews were conducted with a subgroup of EFL teachers and students. The observation of online language learning sessions was also conducted to gain a deeper understanding of the impact of AI-powered collaborative and interactive language learning on EFL teaching in the post-COVID-19 era.

Descriptive and inferential statistics were used to analyse the quantitative data collected from the surveys. In contrast, thematic Content Analysis (TCA) based on Anderson's (2007) 6-step procedure was used to identify key themes and patterns in qualitative data collected through semi-structured interviews and online class observations. The 6-step procedure involves ***Familiarising*** with data, ***Identifying*** themes, ***Coding data***, ***Charting codes***, ***Interpreting findings***, and ***Verifying validity***.

The data were analysed statistically by first tabulating the scores of the post-test subjects as descriptive statistics (mean and standard deviations) and comparing them to the pre-test scores. The data's normality was assessed using the Kolmogorov-Smirnov and Shapiro-Wilk tests. The homogeneity of the data was confirmed through a test of Homogeneity of Variance, and inferential statistics (T-test) were used to determine whether the difference between the two groups was statistically significant, assuming normal distributions. These tests were conducted using SPSS 26.

2.4. Results and Discussion

The findings of this study provide insights into the impact of AI-supported collaborative and interactive language learning on EFL teaching and its implications for EFL teachers and students in the post-pandemic era. However, to fully evaluate and validate them and ensure that they have verified the assumptions made, they should be discussed thoroughly.

2.4.1. Results

The study's results can inform the design and implementation of AI-powered language learning systems in EFL instruction and contribute to the body of knowledge in the AI-powered collaborative and interactive language learning field.

2.4.1.1. Results of the Pre-Test

The initial step in analysing the data involved testing for normality of the data distribution using the Kolmogorov-Smirnova and Shapiro-Wilk tests. Given the sample size, the *Kolmogorov-Smirnova test* was considered the best option. Therefore, the results showed that the significance level for the experimental and control groups was 0.320 and 0.408, respectively, which are both above the commonly accepted p-value of 0.05. The results (cf. Table 3) of the Kolmogorov-Smirnov test suggest that the data from both the group receiving treatment (experimental group) and the group without treatment (control group) of EFL learners are likely to come from a normal distribution. In contrast, the results of the Shapiro-Wilk test suggest that it is improbable that the data from these groups originated from a normal distribution. This discrepancy could be due to how the two tests measure normality.

Table 3. Kolmogorov-Smirnova Tests of Normality

	ID: EFL Learners	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	**Df**	**Sig.**	Statistic	df	Sig.
Pre-TEST	Experimental Group	**.801**	**429**	**.320**	1.351	429	.001
	Control Group	**.843**	**209**	**.408**	1.172	209	.005
a. Lilliefors Significance Correction							
b. Calculated from data							

The homogeneity of variance test, shown in Table 4, compares the variances between the two groups (experimental and control) to determine if they are equal. The results of the test of homogeneity of variance imply that the variance of the experimental group and control group are equal. This conclusion can be made because the significance level (Sig.) is greater than 0.05 in all four test levels. Having equal variances is an essential assumption in many statistical tests, so this result is useful in ensuring that subsequent statistical tests are conducted appropriately. In cases where the variances are unequal, the results' reliability and validity may be at risk, and thus, a different statistical test may be required.

Table 4. Results of the test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
Pre-TEST	Based on Mean	2.347	1	675	.612
	Based on Median	1.081	1	675	1.28
	Based on Median and with adjusted df	1.081	1	667.478	1.28
	Based on trimmed mean	1.629	1	396	.736

Since the pre-test data exhibited a normal and homogeneous distribution, a t-test was deemed appropriate to determine if any observed differences were statistically significant.

Table 4. Results of the test of Homogeneity of VarianceTable 5. Results of three different Chi-Square Tests

	Value	df	Asymptotic significance (2-sided)	Based on these results, it appears that there is no significant association between the variables of the experimental and control groups.
Pearson Chi-Square	5.741[a]	4	.219	
Likelihood Ratio	5.728	4	.220	
Linear-by-Linear Association	3.593	1	.058	
N of Valid Cases	690			
a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 13.47.				

The Pearson Chi-Square test has a value of 5.741 and a degree of freedom (df) of 4. The asymptotic significance (2-sided) is .219 > 0.05, which means there is a 22.1% chance that the test results could have occurred by chance.

The results in table 6 suggest that the majority of respondents (71.2% to 71.7%) are "not at all confident" in their language skills across all skills, except for grammar, where the majority of respondents (54.5%) are "somewhat confident." In grammar, there is a gradually increasing trend in confidence levels, with 12.5% of respondents reporting "not at all confident," 19.3% "not very confident," 20.2% "somewhat confident," 26.9% "fairly confident," and 16.5% "strongly confident." In Personal Communication, there is a similar trend, with 14.1% of respondents reporting "not at all confident," 20.2% "not very confident," 20% "somewhat confident," 27.8% "fairly confident," and 13.7% "strongly confident." However, there is no clear trend in the other skills, and most respondents consistently report being "not at all confident."

Table 6. Descriptive Statistics analysis of different variables related to language skills

	N	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Listening	430	1.29	.468	.219	1.292	.118	1.304	.235
Speaking	430	1.29	.469	.220	1.277	.118	1.252	.235
Reading	430	1.29	.465	.216	1.064	.118	-.493	.235
Writing	430	1.29	.459	.210	1.010	.118	-.778	.235
Pronunciation	430	1.29	.458	.209	1.023	.118	-.749	.235
Vocabulary	430	1.30	.492	.242	1.530	.118	2.890	.235
Grammar	411	3.16	1.295	1.678	-.191	.120	-1.089	.240
Personal Communication	410	3.09	1.282	1.643	-.153	.121	-1.097	.240
Presentation Skills	409	3.43	1.350	1.824	.124	.121	2.858	.241
Listening and Note-Taking	408	3.34	1.300	1.690	-.394	.121	-.954	.241
Improvement Rate	272	3.04	1.378	1.899	-.115	.185	-1.189	.368
Valid N (listwise)	272							

It is important to note that the "Valid N (listwise)" is only 272, meaning that the statistics for the variable 'Improvement Rate' it is representative of the entire population at 63.25 %. This situation is due to non-response bias. Some participants found it challenging to respond to the question – On a scale of 1-5, with 1 being "not at all" and 5 being "significantly," **how much do you feel your English language level has improved in the last twelve weeks?** And this leads to missing data and a smaller sample size for this item. The effects of the missing data for this question on the results are minimal as it was only intended to be a comparative analysis to the other questions, which are related in some manner.

It is worth mentioning that the results of the statistical data analysis conducted on the control group mirrored the findings obtained from the experimental group. The consistency of these findings was confirmed through cross-referencing with qualitative data gathered from a thematic content analysis of both the experimental and control groups. Similarly, the results from the Cambridge test, as shown below (cf. Table 7), align with the previously presented findings.

Table 7. Distribution of language proficiency levels as determined by the CEFR

		Frequency	Percent	Valid Percent	Cumulative Percent
EFL Learners from ExpG	A1 CEFR	286	66.4	66.7	66.7
	A2 CEFR	111	25.8	25.9	92.5
	A2+CEFR	26	6.0	6.1	98.6
	B1 CEFR	6	1.4	1.4	100.0
	Total	429	99.5	100.0	
Missing	System	2	.5		
Total		431	100.0		
ExpG= Experimental Group					

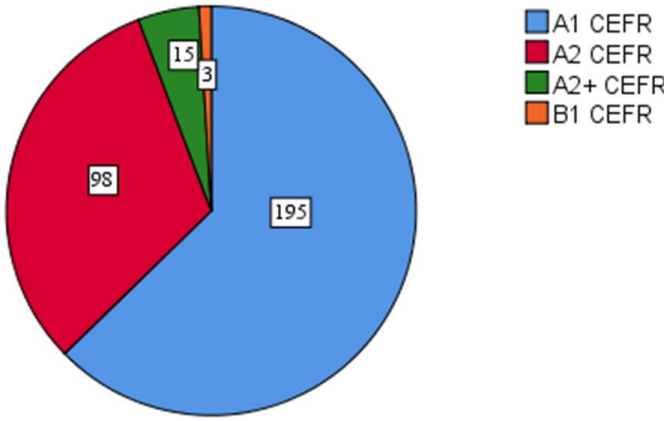

Figure 1. Control Group's EFL learners CEFR-based English language proficiency level

Table 7 and figure 1 present the distribution of Common European Framework of Reference for Languages (CEFR) levels among samples of 431 (Experimental Group) and 311 (Control Group) EFL learners from the English Department and CE.BE.LA.E institutions of the University of Abomey-Calavi in Benin. The CEFR levels used in this study are A1, A2, A2+, and B1. Based on the table and the figure, we can see that a majority of the participants [66.4% (ExpG), 62.7% (ConG)] scored in the A1 CEFR level. About 25.8% (ExpG) and 31.5% (ConG) of the participants scored in the A2 CEFR level, 6.0% (ExpG) and 4.8% (ConG) in the A2+ CEFR level, and only 1.4% (ExpG) and 1% (ConG) of the participants scored in the B1 CEFR level. Overall, the results show that most of the participants have a lower level of proficiency in the English language.

Therefore, it can be concluded that the respondents generally have a low level of confidence in their language skills, except for grammar and personal communication, where there is a moderate level of confidence. The results imply that the respondents may need further training and support to improve their language skills.

2.4.1.2. Results of the Post-Test

The results of the Kolmogorov-Smirnov showed the normality of the distribution of the post-test data. According to the Kolmogorov-Smirnov test, the experimental and control groups had a significance level of 0.510 and 0.302, respectively. Both values are above the commonly accepted p-value of 0.05. These results (cf. Table 8) suggest that the data from both the experimental and control groups of EFL learners are likely to come from a normal distribution. Moreover, the results of the test of homogeneity of variance on the post-test (cf. Table 9) suggest that the dispersion of scores in the two groups is similar, indicating that the variance in the scores is homogeneous between the two groups.

Table 8. Kolmogorov-Smirnova Tests of Normality

Post-TEST	ID : EFL Learners	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	**Df**	**Sig.**	Statistic	df	Sig.
	Experimental Group	**.921**	**386**	**.510**	.591	386	.002
	Control Group	**.994**	**296**	**.302**	.652	296	.007
a. Lilliefors Significance Correction							
b. Calculated from data							

Table 9. Results of the test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.	
Post-TEST	Based on Mean	2.047	1	707	.612	Based on the results, the p-values for all four methods are greater than 0.05, which suggests that there is no significant difference between the variances of the two groups on the pre-test scores.
	Based on Median	1.001	1	707	1.08	
	Based on Median and with adjusted df	1.001	1	767.478	1.08	
	Based on trimmed mean	1.329	1	707	.636	

Table 10 summarises the results of three different Chi-Square Tests, which are used to test the independence between two categorical variables.

Table 10. Results of three different Chi-Square Tests

	Value	df	Asymptotic significance (2-sided)	
Pearson Chi-Square	8.741[a]	4	.519	There is not enough evidence to support the idea that there is a relationship between the Experimental and control group variables.
Likelihood Ratio	8.728	4	.069	
Linear-by-Linear Association	2.593	1	.358	
N of Valid Cases	707			
a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 6.89.				

Based on these results, it can be concluded that there is no statistically significant relationship between the variables of the experimental and control groups.

Table 11. Statistics of AI-powered English language learning experience from 431 respondents.

		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16
N	Valid	431	431	431	431	431	431	428	429	428	428	429	431	428	429	428	429
	Missing	0	0	0	0	0	0	3	2	3	3	2	0	3	2	3	2
Mean		2.49	2.51	2.46	2.50	2.46	2.22	3.28	3.20	3.51	3.44	2.96	2.22	3.28	3.20	3.51	2.96
Median		2.00	2.00	2.00	2.00	2.00	2.00	3.00	3.00	4.00	4.00	3.00	2.00	3.00	3.00	4.00	3.00
Std. Deviation		.553	.574	.508	.562	.508	.840	1.153	1.131	1.107	1.137	.891	.840	1.153	1.131	1.107	.891
Variance		.306	.330	.258	.316	.258	.706	1.330	1.279	1.225	1.292	.795	.706	1.330	1.279	1.225	.795
Skewness		.354	.450	.066	.391	.066	.303	-.107	-.084	-.415	-.269	.544	.303	-.107	-.084	-.415	.544
Std. Error of Skewness		.118	.118	.118	.118	.118	.118	.118	.118	.118	.118	.118	.118	.118	.118	.118	.118
Kurtosis		-.822	-.618	-1.734	-.731	-1.734	-.232	-.936	-.896	-.631	-.850	.108	-.232	-.936	-.896	-.631	.108
Std. Error of Kurtosis		.235	.235	.235	.235	.235	.235	.235	.235	.235	.235	.235	.235	.235	.235	.235	.235
Maximum		4	4	3	4	3	5	5	5	5	5	5	5	5	5	5	5
Sum		1075	1081	1059	1078	1059	956	1403	1374	1503	1474	1268	956	1403	1374	1503	1268

The survey consisted of 16 semi-structured questions about various aspects of language learning and the impact of these AI-powered Google Classroom and IBM Watson tools on the language skills of EFL learners. The mean and median values for each question indicate the average and the middle value of the responses, respectively. The standard deviation and variance values provide information about the spread of the data and how far the individual responses are from the mean. The skewness and kurtosis values measure the symmetry and peakedness of the data distribution, respectively. The standard error of skewness and kurtosis give an estimate of the accuracy of the skewness and kurtosis estimates.

The overall mean score for the survey was 2.49 out of 5, with a standard deviation of 0.55. The median score was 2 out of 5. The results indicate that the respondents had mixed feelings about the effectiveness of these AI-powered tools in assisting with English language learning. The scores for the questions regarding writing and speaking skills were higher (mean scores of 3.28 and 3.20, respectively) compared to the scores for the questions regarding reading and listening skills (mean scores of 2.22 and 2.96, respectively). The results suggest that the AI-powered Google Classroom and IBM Watson tools have helped improve writing and speaking skills to some extent but have not significantly impacted reading and listening skills. &²It is worth noting that the maximum score was 5 and the minimum was 1, which means that some respondents found these tools to be beneficial, while others found them to be not helpful at all. It would be valuable to compare them with the results of the CEFR-based Cambridge proficiency test that the

respondents took after 12 weeks of experience to gain a deeper understanding of these results. This comparison would provide a more complete picture of the language proficiency levels of the participants and how they changed over time with the use of AI-powered tools. Additionally, it would help to understand the effectiveness of these tools in improving language proficiency and to determine if the improvements are aligned with the CEFR levels.

Table 12. Distribution of language proficiency levels as determined by the CEFR

		Frequency	Percent	Valid Percent	Cumulative Percent
EFL Learners from ExpG	A1 CEFR	3	.7	.7	.7
	A2 CEFR	226	52.4	52.4	53.1
	A2+CEFR	145	33.6	33.6	86.8
	B1 CEFR	57	13.2	13.2	100.0
Total		431	100.0	100.0	

ExpG= Experimental Group

This table shows the distribution of CEFR (Common European Framework of Reference for Languages) levels among English as a Foreign Language (EFL) learners in the experimental group (ExpG). The four CEFR levels listed are A1, A2, A2+, and B1. The "Frequency" column shows the number of participants at each level, with 226 participants at the A2 CEFR level being the largest group. The "Percent" column shows the percentage of participants at each level based on the total number of participants (431). The "Valid Percent" column shows the percentage of participants at each level based on the number of valid responses (431). The "Cumulative Percent" column shows the cumulative percentage of participants at each level and all levels below it.

As per the findings displayed in Table 12 and Figure 2, the A2 CEFR level was found to be the most frequent among the participants (52.4%) in the experimental group, followed by the A2+ CEFR level (33.6%). A small number of participants were at the A1 CEFR level (7%) and B1 CEFR level (13.2%). The findings suggest that most participants in the experimental group had a proficiency level at the A2 CEFR level or higher, indicating a relatively high level of language proficiency. Additionally, the results reveal that the participants in the experimental group were evenly distributed across various CEFR levels, indicating a diverse range of language proficiency within the group.

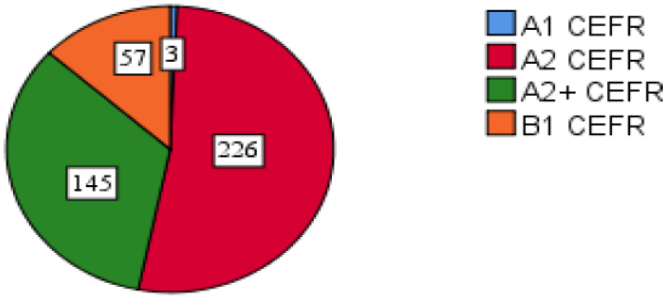

Figure 2. Experimental Group's EFL Learners CEFR-based English Language Proficiency Level

On a scale of 1-5, with 1 being "not effective at all" and 5 being "very effective", "how would you rate the effectiveness of Google Classroom AI compared to IBM Watson Natural Language Understanding in helping EFL learners improve their language skills

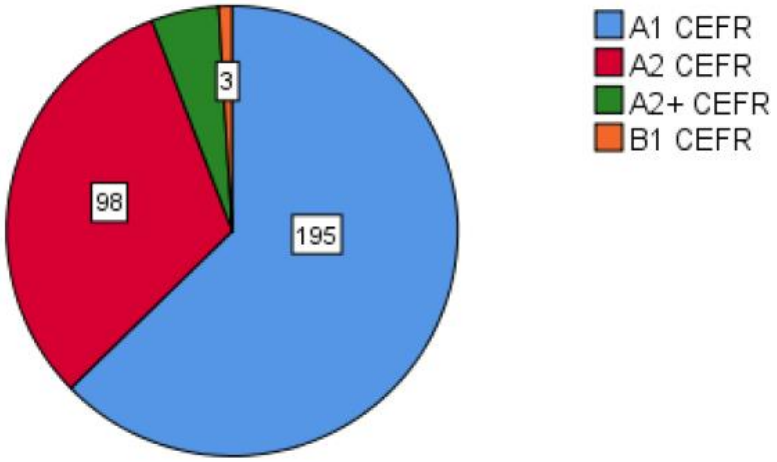

Figure 3. Control Group's EFL Learners CEFR-based English Language Proficiency Level

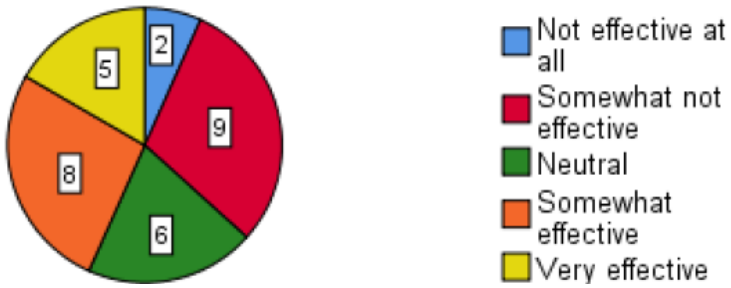

Figure 4. Statistics of EFL teachers answers to the impact of AI - Powered tools on EFL learners

By comparing the results of both groups (experimental and control), it can be observed that a higher percentage of participants in the experimental group were at the A2+ and B1 levels (33.6% and 13.2%, respectively) compared to the control group, which has a higher percentage of participants at the A1 level (66.7%). The control group has a lower percentage of participants at the A2+ and B1 levels (6.1% and 1.4%, respectively) than the experimental group. This difference in the distribution of CEFR levels between the participants of the two groups (ExpG and ConG) suggests that the use of AI-powered tools has positively impacted the English language proficiency of the Beninese EFL learners in the experimental group. It suggests that the AI-powered tools may have helped the experimental group improve their English language proficiency. It is possible to interpret this observation as an indication of the positive impact of AI-powered tools on English language learning. Additionally, the opinions gathered from the 30 teachers involved in this study align with previous findings. However, for some of them, drawing a definite conclusion about the impact of AI-powered tools without further data and analysis remains inconclusive.

Based on the results of the qualitative data analysis obtained from the respondents, it appears that a definitive conclusion on which of the two AI-powered tools, Google Classroom AI and IBM Watson Natural Language Understanding, is better cannot be drawn at this time. Further data

collection, analysis, and observation over a more extended period are necessary to fully assess each tool's effectiveness and determine which is better.

2.4.2. Discussion

The COVID-19 pandemic has brought unprecedented challenges to the field of education, particularly in teaching English as a foreign language (EFL). The traditional classroom-based approach to language learning has been disrupted, leading to a rapid shift toward online learning. In this context, AI-powered collaborative and interactive language learning is becoming increasingly relevant to enhance EFL teaching in the post-COVID-19 era. This section discusses the study's results that examined the impact of *AI-powered Google Classroom* and *IBM Watson Natural Language Understanding tools* on Beninese EFL teaching in the post-pandemic era.

The study results showed that AI-powered collaborative and interactive language learning could enhance EFL teaching in the post-pandemic online environment. The results reveal a mixed response from the respondents about the efficiency of AI tools in English language learning. The average score was 2.49 out of 5, with a standard deviation of 0.55 and a median score of 2 out of 5. However, this study demonstrates that implementing AI-assisted collaborative online learning can enhance engagement in English as a Foreign Language (EFL) instruction, improve learning outcomes for EFL students, and increase teacher satisfaction.

They suggested that the *AI-powered Google Classroom* and *IBM Watson tools* have helped improve writing and speaking skills to some extent but did not significantly impact reading and listening skills. However, this study found that the results from the Cambridge Proficiency post-test showed high scores at the A2 CEFR level (52.4%), A2+ CEFR level (33.6%), and minimal scores at the A1 CEFR level (7%), and B1 CEFR level (13.2%). This observation is in contrast to the results from the pre-test, where the majority of participants were at the A1 CEFR level (66.4%), and a lower percentage were at the A2 CEFR level (25.8%) and A2+ CEFR level (6.0%), with only a small number at the B1 CEFR level (1.4%). The overall results indicate that most participants who participated in the study and used the two AI-powered collaborative and interactive language learning tools had a relatively high level of proficiency, with most of them at or above the A2 CEFR level. They suggest that the use of AI-powered tools has had a positive impact on the English language proficiency of the Beninese EFL learners. These findings align with prior research that supports the notion that AI-powered collaborative and interactive language learning can enhance English as a Foreign Language (EFL) instruction in the post-pandemic online setting (Chen, Chen, & Lin, 2020). Therefore, the first research assumption is proven to be valid.

Furthermore, the results obtained through online class observations support the notion that using AI-assisted collaborative e-learning can lead to enhanced student engagement in EFL teaching, improved learning outcomes, and increased teacher satisfaction. They align with the findings of Huang, Lu, and Yang (2023) and reinforce the study's second hypothesis.
In addition, the qualitative data gathered from 30 EFL instructors highlight the significance of enhancing teachers' proficiency in utilising Information and Communication Technology (ICT) to integrate AI-assisted collaborative e-learning into EFL teaching successfully. The results suggest that such professional development can equip teachers with the necessary skills to effectively use AI-enabled tools in their EFL teaching. This finding supports the third research hypothesis (Hennessy et al., 2021).

The results of this study offer valuable insights into the influence of AI-backed collaborative and interactive language learning on English as a Foreign Language (EFL) education and its

implications for both EFL teachers and students in the aftermath of the pandemic. These findings can help inform and guide future EFL teaching and learning practices.

To sum up, the effectiveness of AI-powered tools in English as a foreign language (EFL) learning is still a topic of ongoing research and debate. However, substantial evidence indicates that EFL learners can derive multiple advantages from AI-powered tools, such as tailored feedback, real-time error correction, and gamification features that can enhance the level of engagement and enjoyment associated with language learning. Furthermore, AI-powered tools are well-suited for EFL instruction due to their flexibility and convenience, enabling learners to access an extensive array of learning materials and resources.

It is crucial to remember that AI-driven tools are not intended to substitute conventional language teaching methodologies; they are intended to complement them. The most effective EFL learning programs likely involve human interaction and technology-based tools. Additionally, it is important to ensure that AI-powered tools are evaluated regularly and that their results are validated by independent research to ensure their continued effectiveness and that they are being used to support the best outcomes for language learners.

3. CONCLUSIONS

The COVID-19 pandemic has significantly changed traditional language learning, necessitating a transition to online teaching and exploring new approaches to language education. This study, conducted at the University of Abomey-Calavi in Benin, aimed to assess the effectiveness of incorporating AI-powered collaborative and interactive methods, such as the use of *AI-based Google Classroom* and *IBM Watson Natural Language Understanding tools*, in teaching English as a foreign language (EFL) within a post-pandemic online learning context. This mixed-methods research design, which involved surveys, semi-structured interviews, and observation of online language classes, offered a comprehensive perspective on the challenges and opportunities presented by AI in EFL teaching. The study's outcomes provide insight into the most efficient teaching methods and evaluation techniques, the valuable support provided by technology, and the significance of teacher professional development for integrating AI-assisted collaborative e-learning in EFL instruction.

Moreover, the research also reveals the benefits and drawbacks of AI-based tools, specifically Google Classroom and IBM Watson Natural Language Understanding, implemented in the current research context. The benefits include the ability of AI algorithms to personalise learning, automate grading and feedback, enhance accessibility in education through real-time translation, text-to-speech, and other accessibility features, and increase engagement with interactive and engaging learning experiences. However, the limitations of these AI-powered tools include potential biases and limitations related to gender, race, or culture, reliance on human input, technical issues such as compatibility with other software, slow response times, and limited customisation options, as well as a reduced need for human interaction.

It is crucial to acknowledge the limitations of AI in language education and the importance of human interaction and feedback. Although AI-powered tools have the potential to automate some aspects of language learning, such as grading and feedback, they should not be seen as a substitute for human interaction and personalised instruction. The best results in language education are likely to be achieved by using AI-powered tools should as a complement to, not a replacement for, human interaction and feedback in language education. This approach will help ensure that language learners receive the best possible education, taking advantage of both human interaction's strengths and technology's benefits.

In sum, the study sheds light on the potential of AI-assisted interactive and collaborative language learning for EFL teaching and offers valuable insights for EFL students and teachers to navigate the changing landscape of online language education. Based on the findings, the use of AI technology has the potential to enhance and improve language education, especially in the post-pandemic era.

To build on the current research and make meaningful contributions to the field of EFL learning and AI, future researchers should:

- Extend the scope of research by including a broader range of AI-powered tools and a more diverse sample of language learners from different backgrounds, cultures, and proficiency levels.
- Investigate the impact of AI-powered tools on different aspects of EFL learning, including language acquisition, motivation, engagement, and language proficiency.
- Use mixed-methods research designs, combining both quantitative and qualitative data collection and analysis techniques, to obtain a more thorough comprehension of the influence of AI on EFL learning.
- Consider the ethical implications of using AI in language education, such as data privacy, data security, and algorithmic bias, and take steps to mitigate these risks.
- Foster collaboration between educators, researchers, and technology developers to explore the best ways to integrate AI-powered tools into EFL learning and to maximise their impact on language education.
- Evaluate the long-term impact of AI on EFL learning and its potential to transform language education in the future.

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to everyone who played a part in completing this research paper, with special recognition for the teacher colleagues and EFL students of the English Department and the Beninese Center for Foreign Language (CE.BE.LA.E) of the University of Abomey-Calavi. Your support was instrumental in the success of this project, and I am sincerely thankful for your valuable contributions.
Sincerely,

REFERENCES

[1] Abalkheel, A. (2022). Amalgamating Bloom's taxonomy and artificial intelligence to face the challenges of online EFL learning amid post-COVID-19 in Saudi Arabia. *International Journal of English Language and Literature Studies*, *11*(1), 16-30.

[2] Abd-Alrazaq, A., Alajlani, M., Alhuwail, D., Schneider, J., Al-Kuwari, S., Shah, Z., ... & Househ, M. (2020). Artificial intelligence in the fight against COVID-19: scoping review. *Journal of medical Internet research*, *22*(12), e20756.

[3] Anderson, R. (2007). Thematic content analysis (TCA). *Descriptive presentation of qualitative data*, 1-4.

[4] Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *Ieee Access*, *8*, 75264-75278.

[5] Dobrev, D. (2012). A definition of artificial intelligence. arXiv preprint arXiv:1210.1568.

[6] Goldie, J. G. S. (2016). Connectivism: A knowledge learning theory for the digital age? *Medical teacher*, *38*(10), 1064-1069.

[7] Hennessy, S., D'Angelo, S., McIntyre, N., Koomar, S., Kreimeia, A., Cao, L., ... & Zubairi, A. (2021). Technology, teacher professional development and low-and middle-income countries: technical report on systematic mapping review. [Policy Brief]. EdTech Hub. https://doi.org/10.53832/edtechhub.0080. Available at https://docs.edtechhub.org/lib/7S9CUP77

[8] Holmes, W., Persson, J., Chounta, I. A., Wasson, B., & Dimitrova, V. (2022). *Artificial intelligence and education: A critical view through the lens of human rights, democracy and the rule of law*. Council of Europe.

[9] Huang, A. Y., Lu, O. H., & Yang, S. J. (2023). Effects of artificial Intelligence–Enabled personalised recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers & Education*, *194*, 104684.

[10] Jiang, R. (2022). How does artificial intelligence empower EFL teaching and learning nowadays? A review on artificial intelligence in the EFL context. *Frontiers in Psychology*, *13*.

[11] Junaidi, J. (2020). Artificial intelligence in EFL context: rising students' speaking performance with Lyra virtual assistance. *International Journal of Advanced Science and Technology Rehabilitation*, *29*(5), 6735-6741.

[12] Kalina, C., & Powell, K. C. (2009). Cognitive and social constructivism: Developing tools for an effective classroom. *Education*, *130*(2), 241-250.

[13] Lee, J., Kim, M., & Park, H. (2022). The impact of AI-powered collaborative language learning on enhancing EFL teaching in the post-COVID-19 era: A social constructivist perspective. *Journal of Educational Technology Development and Exchange*, *5*(1), 1-10.

[14] McArthur, D., Lewis, M., & Bishary, M. (2005). The roles of artificial intelligence in education: current progress and future prospects. *Journal of Educational Technology*, *1*(4), 42-80.

[15] MND. (2023). Le Bénin se dote d'une stratégie nationale d'intelligence artificielle et des mégadonnées. [Retrieved on 01-28-2023 from https://numerique.gouv.bj/publications/actualites/le-benin-se-dote-d-une-strategie-nationale-d-intelligence-artificielle-et-des-megadonnees#]

[16] Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, *2*, 100020.

[17] Pedro, F., Subosa, M., Rivas, A., & Valverde, P. (2019). Artificial intelligence in education: Challenges and opportunities for sustainable development.

[18] Piaget, J. (1953). *The origins of intelligence in children*. New York, NY: Basic Books.

[19] Pokrivčáková, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education*.

[20] Schmaus, R. (2022, 1er dec). Intelligence Artificielle : les définitions à connaitre pour briller en société [Consulté le 30 Janvier 2023 à partir de https://www.talkwalker.com/fr/blog/intelligence-artificielle]

[21] Siemens, G. E. O. R. G. E. (2005). Connectivism: A learning theory for the digital age. International Journal of Instructional Technology and Distance Learning. *Online] retrieved from: http://www. idtl. org/Journal/Jam _05/article01. html*.

[22] Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, *12*(2), 257-285.

[23] Sweller, J., Ayres, P., Kalyuga, S. (2011). Cognitive Load Theory in Perspective. *Cognitive Load Theory*, 237-242.

[24] Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational psychology review*, 251-296.

[25] Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard university press.

[26] Vygotsky, L.S. (1962). *Thought and language*. Cambridge, MA: MIT Press (original work published in 1934).

[27] Warschauer, M. (1996). Computer-assisted language learning: An introduction. *Multimedia language teaching*, *320*.

[28] Warschauer, M. (1997). Computer-assisted language learning: An introduction. In M. Warschauer & R. Kern (Eds.), Network-based language teaching: Concepts and practice (pp. 1-12). Cambridge University Press.

APPENDIX

Questions in the 5-Point Likert Scale Semi-Structured Questionnaire – (Q11 to Q16 were intended for EFL teachers and Learners)

Q1 How effective do you feel Google Classroom AI is in assisting with your English language learning?

Q2 To what extent do you feel IBM Watson's Natural Language Understanding capabilities improve your understanding of English texts and writing?

Q3 How helpful is Google Classroom AI in helping you practice speaking English?

Q4 How effective do you believe IBM Watson's Natural Language Understanding is in improving your writing skills in English?

Q5 To what extent have the Google Classroom AI and IBM Watson helped you become more confident in using English in everyday communication?

Q6 To what extent have your listening skills improved since using the AI-powered Google Classroom and IBM Watson tools?

Q7 To what extent have your writing skills improved since using the AI-powered Google Classroom and IBM Watson tools?

Q8 How much have your speaking skills improved due to using the AI-powered Google Classroom and IBM Watson tools?

Q9 How effective do you feel the AI-powered Google Classroom and IBM Watson tools have improved your reading skills in English?

Q10 To what extent have the AI-powered Google Classroom, and IBM Watson tools helped you better understand English grammar and vocabulary?

Q11 How effectively do you believe the Google Classroom AI and IBM Watson Natural Language Understanding tools assist EFL learners with their language skills development?

Q12 In your experience, to what extent have the AI-powered Google Classroom and IBM Watson tools helped EFL learners improve their listening skills?

Q13 How helpful are the Google Classroom AI and IBM Watson tools in improving EFL learners' speaking skills? (

Q14 To what extent do AI-powered Google Classroom and IBM Watson tools help EFL learners improve their reading and writing skills?

Q15 How would you rate the overall impact of the AI-powered Google Classroom and IBM Watson tools on the language development of EFL learners?

Q16 On a scale of 1-5, with 1 being "not effective at all" and 5 being "very effective," how would you rate the effectiveness of Google Classroom AI compared to IBM Watson Natural Language Understanding in helping EFL learners improve their language skills (Please note that this question was optional for EFL learners but essential for EFL teachers to consider.)

Q17 On a scale of 1-5, with 1 being "not at all" and 5 being "significantly," how much do you feel your English language level has improved in the last twelve weeks?

AUTHOR

Coffi Martinien ZOUNHIN TOBOULA is a Senior Lecturer at the University of Abomey-Calavi's Department of Literature, Languages, Arts, and Communication (FLLAC). His area of expertise is in Linguistics and Didactics of the English language, for which he earned a Ph.D. He is also well-versed in psycho-pedagogy, learner psychology, anxiety management, oral communication promotion, leadership, extracurricular activities, and innovative teaching and research techniques in Teaching English as a Foreign Language (TEFL). He is a skilled communicator and a member of Toastmasters International. Moreover, he actively promotes and is a founding member of various English language learning communities, including the Universal Toastmasters Club of Cotonou. Dr. TOBOULA works as a freelance translator and interpreter, manages a language laboratory, and is passionate about writing and research in education and human psychology.