



Effectiveness of AI Tools in English for Occupational Purposes: Enhancing Pragmatic Competence and Fluency

Humaira Nadeem

Department of English language and linguistics studies, University of Sargodha

Hamna Qaiser

Department of English language and linguistics studies, University of Sargodha

Sana Nawaz (corresponding author)

Department of English language and linguistics studies, University of Sargodha

sana.nawaz@uos.edu.pk

Abstract

This present study aims to determine the efficacy of the use of the AI-generated conversational chatbots in the pragmatic competence and genre-precise fluency of English for Occupational Purposes (EOP) students who need outcome-driven and contextually appropriate assessment. Mixed methods were used when evaluating 40 undergraduate business students, with the experimental group in an email imitating setting with a custom AI Chatbots, while the control group were given with the feedback of moderator. To qualify the politeness strategies, lexical and structural changes towards the professional genres, both the pre test and post test and as well as qualitative reflections were analyzed. The study shows that in achieving the pragmatic competence knowledge and to decrease the anxiety frequent communications with AI chatbots work. Overall, the use of Artificial Intelligence in specialized language courses develop independence and productively connects teaching and learning with professional communication. The results of this study would be helpful inputs for those who would like to teach language and want to design the syllabus productively with the help of developing technologies in the context of EOP.

Introduction

The appearance and rapid advancement of AI has revolutionized the learning ecosystem especially in specialized areas such as English language learning. In modern applied linguistics, the both English for Occupational Purposes and English for Academic Purposes require the students to have a genres knowledge, precision of language and a high communicative competence. The physical setting in which learners use these language context are often challenging and difficult to meet with traditional teaching approaches that fail to offer personalized and real time reactions to learners. AI assistants and virtual assistants like conversational agents (CAs) have thus become emerged to fill the gap between the classroom application and real life activities. The application of technology for language learning started with the passive component, language acquisition but contemporary applications of artificial intelligence support students in engaging actively, by making them part of the activity, having them interact with actual professionals and receive individualized language input (**Godwin-**

Jones, 2022; Oleksandrivna, 2026).

The direct impacts of these AI chatbots and interactive assistants on the targeted learning consequences is a substantial area of need. Few studies have looked at errors from the perspective of semantic content, apart from a few basic grammatical errors which can be handled by language-inputting applications such as Grammarly. But this is not enough for correct grammatical communication, however: it needs to be aware of situationally and professionally distinct communication, structured and must be full of politeness tactics. As argued by **(Pei, 2025; Zehetgruber, 2025)**. EOP students need to learn certain types with a particular nuance, as in the case of executive summaries, formal emails and project reports, one might sail through the whole communication, only to be confronted by a knock-on the head and a completely new tone. This need underscores the importance of conversational AI tools, which can emulate the conversational and interactive nature of human communication, and creates an opportunity to such an experiential learning environment that one wouldn't have otherwise.

The purpose of this research study is to investigate the influence of systematic discussion with artificial intelligence (AI) conversational chatbots on literacy proficiency and pragmatic development skills of university students specializing in business. This research is intended specifically at the EOP context to find out if interactive real time and automated response can publicize the real accommodations in the linguistics at a thick level that suit to the professional space. The central inquiry is for learners to gain situational fluency, or using language to suit any given occupation's dynamic and purpose. This research considers AI as an adaptive and scalable tool to offer repeated, intensive practice needed for mastery instead of a substitute for man instructors. **(Zehetgruber, 2025)**.

As argued by **(Fan and Ma (2022) and Zulkornain et al. (2023)**, this research holds great importance as it aims to provide an empirical but structured evaluation of the use of AI in a target-oriented curriculum. The Research investigates both quantitative and qualitative indicators of writing scores and learner confidence in writing to address this key gap in the existing EOP pedagogies of writing. Class size and the time limits can be a problem for instructors, and they may not be able to describe in detail what they did well and what they need to work on for each draft. Further exploration of the ability of AI tools to do routine formative assessments has the potential to enable teachers to invest more time in conceptually guiding students in the classroom. In the end, the goal of this project is to create a comprehensive, well-balanced model to use AI when equipping language students with the skills necessary to undertake effective communication in the contemporary professional environment.

Literature Review

The evolution of computer Assisted language learning is the Mobile Assisted language learning. The integration of computer Assisted language learning has been developed into the language learning that is Mobile Assisted language learning. As discussed by **(Godwin-Jones (2022))**, initially, digitization was seen as a mechanism primarily to enhance the appearance of a dictionary, or to serve it the same manner as a structural drill machine, with repetition of behavior. But with the advent of complex elegantly designed big language models, it is more applicable to interactive, communicative competence. For any EAP or EOP contexts where learners need specific, context specific skills, the use of technology should be measure by scaffolding complex skills. Previous studies suggest that it is not enough for specialized language learners to learn a vocabulary; Previous studies suggest that for a language learners just vocabulary learning is not enough; what they need is rather, to understand and memorize the meanings of the vocabulary in the specific situations of certain genres like socio-professional genres.

This study employs a number of existing Second Language Acquisition (SLA) frameworks to facilitate better understanding of role that AI tools can play in language learning. First, language acquisition is related to language exposure in the form of comprehensible input that is a little above the learner's level. **(Krashen, 1982, 1985)**. First, according to Krashen's Input Hypothesis, language acquisition results in "grasping" a little more than you can, because that is exactly as much as it needs" to be acquired. The generative dialog systems are better at doing this because they will adapt their vocabulary and structure to match the input from users, building a speech community that is unique to each individual. Secondly, Swain's Output Hypothesis states that in order for the learner to process language, he must do it critically when he faces the communicative failures. According to **Swain (1985, 1995)**, The learner must identify the incomplete knowledge or competence when s/he is provided with a hint by the AI chatbots that s/he has not understood the prompt; s/he then needs to adjust his/her output for a successful communication. In addition, Schmidt's Noticing Hypothesis focuses on the fact that input must be 'noticed' in order to be 'kept' and thus incorporated by the learner. **(Schmidt, 1990, 2001)**. **As observed by (Tian et al. (2025) and Damanik (2025)**, Feedback in the traditional sense (via writers' notebooks, for instance) of a writing class is often delivered days or even weeks after the writing is complete, meaning that students are separated from their initial decisions about their thoughts for writing. AI can answer this by providing feedback to learners immediately as they are writing. If an automated system makes an imprecise signal to the student when the error appears, When the error appears and the automated system gives an inaccurate signal to the learner, one would expect that the learner would also see it as soon as it comes and will be perplexed by seeing reason behind the change, and will be able to integrate it into his/her working memory.

For automated writing evaluation the existing tools have some positive and negative that have been taken from previous studies. Many software programs that emphasize syntax and orthography will assist learners to remove the superficial errors, yet there may be little attention to the production of some pragmatic appropriateness **(Taguchi, 2020)**. An EOP language can be grammatically perfectly correct, but totally incongruous because of a confrontational tone and/or inadequate hedging **(Zehetgruber, 2025)**.

. Furthermore, In literature when there is a focused on teaching then the distinction between a grammar checker (passive) and chatbots (active) is not always clear. This study fills these gaps by varying the interactive condition of conversational, artificial intelligence, and offering a clear comparison between the effect of dynamic negotiation of meaning in relation to the process of negotiating meaning in professional writing compared to more traditional pedagogical approaches **(Long, 1996; Tseng & Warschauer, 2023)**.

Methodology

There is a use of mixed methods research and the experimental research design to assess the effectiveness of Chatbots on EOP writing development. By using quantitative data standardized pre tests and post tests were conducted to determine changes in fluency, appropriateness and in language accuracy. Structured introspective journals were used to gather qualitative data using thoughtful ideas, especially from the participants at the end of the intervention. In this way, any quantifiable increases in language production are made to balance with the experience and mental aspects of the learners as they manage within their interactions with the digital landscape.

There were forty undergraduate students who were the participants of the subjects, taking a compulsory course in professional communication. These individuals had similar linguistics background and had same high school, with intermediate levels of English proficiency

that was confirmed by the pre-test placement. There were 20 students and these students were divided in 2 groups the experimental group and one control group and the subjects were selected from the assignment randomly. The sampling procedure was a random selection. The students in the experimental group wrote their assigned tasks that are weekly on a special conversational AI scenario that was set up to interact with students in a manner designed to mimic a workplace scenario, and received real life remarks from the course moderator in the traditional way, as did the students in the control group. While the participants of experimental group used the same specialized conversational interface and wrote the same amount of material each week and that was designed to strive a workplace situation and provided real time reactions in the traditional way by the course moderator, as did the students of the control group.

The content was based on a genre of formal business writing. The teaching course lasted for six continuous weeks. Following the task-based business simulation design established by **Pei (2025) and Zehetgruber (2025)**, Both groups were given a different situation to write a solution to every week with the expectation that they could articulate their proposed solution in writing. Every week both groups were given a new scenario within a company to write a solution to; whether that meant negotiating a delay on a contract, an issue with a client, or a formal request for a change to a budget from a superior. As structured in the comparative design by **Kartika (2024) and Tian et al. (2025)**, The control group composed responses themselves and provided them weekly and got detailed grammatical, layout and toning feedback in writing after seven days. The participants in the experimental group directly talked to the AI chatbot that has been programmed to be the corporate interlocutor regarding the set of scenarios, correcting the statements, suggesting alternative words or giving ratings on tone after each statement turn.

There were four main criteria for the assessment on professional writing: grammatical accuracy, lexical choice, structural coherence, and pragmatic appropriateness. All of the dimensions were graded on a 1 - 5 scale and for each dimension the maximum possible score was 20 on the script. A correlation formula was used to calculate inter raters reliability where there was high correlation between the two evaluators indicating the quantitative scoring process was objective. (**Fan & Ma, 2022; Zulkornain et al., 2023**).

Data Collection

Data was collected throughout the six week academic module, thus comprising three data sets in order to provide a robust foundation for analysis. The scores of the pre-test and the post-test were used as the primary data set. In week one all forty participants attempted a written complaint with project proposal draft in a controlled classroom without access to external digital help. The pre-test in week one involved all four subjects writing a written complaint and draft of a project proposal in a controlled classroom without any external digital support. The post-test was given at the end of week six with exactly the same administration with the exception that the participants were asked to respond to a new set of relevant workplace scenarios of the same level of linguistic complexity.

The second set of data was created from the interactions between the experimental group and their instruments. As the students interacted with the AI chat bot to go back and forth with it in the Digital interface, all the text has been captured, including the initial drafts by the student and the immediate feedback that the AI generated, as well as students' revisions in the chat thread. This automated tracking gave an unedited record of the writing process as students altered their language in different ways as they received real time prompts from the AI, and identified patterns of change in terms of their errors.

The third data set was the qualitative reflective journals from the experimental group that were taken at the end of the study week. Students were given wide-open prompts to

describe their experience in using these for their level of confidence, technical issues, and how immediate feedback from AI integrated with the content supported their understanding versus that of a traditional instructor's notes. Both of these journals were provided electronically and thematized in a text-based database. Qualitative feedback was used to gain an insight into affective filters of the learners to capture their emotional aspects, such as anxiety in writing, motivation change and so on, which quantitative feedback could not capture.

Data Analysis

To assure the baseline level of equality between the groups the quantitative evaluation was started by comparing between scores of both groups. On the control group the mean pre-test scores were 12.45 for 20 points, and 12.60 for 20 points in the experimental group. This minor difference was not significant as determined by a statistical independent t test, and confirmed that both groups had the same business writing skills before the intervention. There was uniformity of baseline that conclude that differences in instructional feedback methods were creating the differences in performance.

After the six weeks, between the two groups significant differences between average literature and cognitive development scores were observed. The scores of this group of students with conventional weekly instructor assessment reveals a slight gain with a mean of 14.10, which is 1.65 points higher than before. In Contrast the experimental group, who used the conversational AI chatbot, with a mean score shows an impressive improvement of 4.25 points after the test of 16.85. By using a pre-test / post-test t-test for the experimental group these results were compared to determine if the results were substantial, and the figures were used to prove that in terms of meaning and statistically, the experimental group improved significantly.

Table 1

The Mean scores from the various assessment methods for the Control and Experimental groups.

Judging Criterion (Out of 5.0)	Control Test	Pre-Test	Control Post-Test	Post-Test	Experimental Pre-Test	Experimental Post-Test
Grammatical Accuracy	3.20		3.65		3.25	4.30
Lexical Selection	3.10		3.50		3.15	4.15
Structural Coherence	3.15		3.55		3.10	4.20
Pragmatic Appropriateness	3.00		3.40		3.10	4.20

A precise analysis of the numerous categories of grades revealed that there were the greatest gains in pragmatic appropriateness and grammatical accuracy. The mean pragmatic score of the AI group went from 3.10 to 4.20, but the mean score for the control group increased from 3.00 to 3.40. The difference shows that real-time feedback on the results of some of the students' stylistic choices related the direct communication of their texts with an interlocutor of an AI is aided students to understand this. The students saw how the same words were used in a different phrase, thus modifying the Chabot's voice, and making the abstract concepts of register and tone a lot more real.

Qualitative data, gathered with the students' analytical journals, was qualitatively analyzed and demonstrate some recurring themes that are related to the experience of the students. The common theme was marked gradual decrease in writing anxiety: 70% of the experimental writers showed a high level of relief making errors in front of the neutral digital interface as compared to the human mentor. A significant point that emerged was the value of timely feedback: students felt that instantaneous evaluation empowered them to grasp and rectify their mistakes while they were still salient and in their line of thinking. However, the remaining 30% of users reported some aggravations related the inflexibility of AI, such as when it misinterpreted what they were trying to convey or gave too preset responses for corrections.

Discussion

More over, both the qualitative and quantitative results confirm that the use of tools in EOP enhance the outcomes of students' writing, when compared with using them in isolation. The experimental group suggests the theory about the value of formative assessment in real time is positively confirmed because experimental group showed the significant gains in post test scores. "I knew that because I have to constantly amend and rerite the texts and adapt them to the context, from the Swain perspective, I can only say that the chatbot forces me to continue modifying the text and produce something more specific and appropriate for the context because it is interactive." This repeated process produces language processing much faster than reviewing graded assignments simply, which is a passive approach.

This pragmatic appropriateness has increased significantly, is what makes it so useful to study interaction in a specialized language acquisition in a contextualized way. The politeness strategies and register must be mastered in professional writing. The logs produced by the interaction revealed that when the chatbot was not able to interpret informal input students began adapting their input. The input shifted to being formal and more standardized. This behaviour is consistent with Schmidt's Noticing Hypothesis which shows that there are pragmatics attending to boundary abstraction that are immediately actable and noticable, depending on the context. The chatbot becomes a non risky, virtual, environment for the learners to practise communication strategies, as if in a workspace, but without major reverberations.

The artificial intelligence interface was learner-friendly with a non-judgemental and neutral tone, which aided participants overcome their emotional or psychological barriers and avoid the "affective filters" that might otherwise have inhibited use of longer sentences and new words. Because the tone of the AI interface was non-judgemental and student-friendly, they were able to overcome their fear of "affective filters" and experiment with longer sentences and new words. This fluency gain is directly evident as frequent repetition can help to structure the pattern into automation and contribute to building confidence longer term.

But, the instructional risks stated from the feedback of qualitative analysis will need to be handled with care. Some students' corrections are drafting, revealing that heavily relying on a single artificial intelligence system might create a generic, unified style and texture of writing, without any sense of voice. Moreover, since AI models are focused on identifying patterns rather than understanding the interpretation of meaning, they may sometimes misinterpret complicate sentences of users, leading to inaccuracies. Moreover, AI models may fail to capture the meaning of creative and valid sentences. However, AI chatbots are not a good substitute for feedback in-depth information, particularly in the culturally responsive manner of a human instructor. Finally, the best EOP curriculum will use artificial intelligence as help for checking and drafting early, while allowing time for teachers to work on complex rhetorical strategies and/or extensive critical analysis.

Conclusion

This project shows the potential of using AI tools for conversational or interpersonal skills to support or help EOP students. These tools capture the student's instant feedback (including text usability quality) and offer realistic interaction to gain Pragmatics awareness, grammatical accuracy and structural fluency, faster than the classroom instruction alone would do. The study supports the capability of generative artificial intelligence to effectively reproduce professional settings, providing larger and smaller, individualized learning practice to fulfill the demands for specialized language learning.

The findings have instant application in the real world or practical world of the curriculum designer and/or the language educator in an institutional context. The presence of AI tools within homework modules enables institutions to provide individualised support to learner's learning without adding to teacher workloads or pressure. The virtual assistants can be used by instructors or educators to run fundamental grammar and formatting checks, leaving class time for intense collaborative activities, which increase their level of rhetorical growth. The study has provided some perceptions however there were some limitations that indicated future research gap. The limited duration of research (six weeks) and number of participants (40 students) interfere the possibility of confirming the outcomes of the research for the long term with more learners or students in broader samples. Furthermore, in this project only written communication in a particular business context was developed. Further study would be advantageous in understanding how the V/VXIV apps for voice-to-text collision on oracy, presentation and 'real time negotiation tools' both in the workplace environment and in the classroom.

Finally, the integration of AI into the learning process and teaching should not be seen as a substitute for traditional teaching methods or a risk to the role of teachers. Digital or computerized communication is still evolving, so it will be crucial to continue to use these interactive methods in the EAP/EOP curriculum to prepared the students to overcome the communication challenges facing the business climate around the world.

References

- Fan, N., & Ma, Y. (2022). The effects of automated writing evaluation (AWE) feedback on students' English writing quality: A systematic literature review. *Language Teaching Research Quarterly**, *28*, 53–73. <https://doi.org/10.32038/ltrq.2022.28.03>
- Fatihah, A. (2026). The role of artificial intelligence (AI) in enhancing English language skills: A systematic review. *International Journal of Pedagogical Studies and Social Science*, 4(1), 88–99.
- Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, 26(2), 5–24.
- Lu, C. T. (2025). An intelligent English-speaking training system using generative AI and speech recognition. *Applied Sciences*, 16(1), 189–204. <https://doi.org/10.3390/app16010189>
- Oleksandrivna, N. O. (2026). AI-enhanced ESP instruction for technical students. *Publishing House Education and Science*, 3, 24–38.
- Pei, W. (2025). Constructing and practicing a precision teaching model for business English writing course empowered by AIGC. *ISRG Journal of Arts, Humanities and Social Sciences*, 3(6), 112–125.
- Zehetgruber, M. (2025). Evaluating AI-powered chatbots interfaces in specialized occupational communication and lexical acquisition. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, 19(38), 45–59.