

Using Voice-Enabled AI Chatbot to Improve Academic Presentation Skills

Aisha Aslam¹  and Saqib Aziz² 

¹The University of Queensland, Australia

²Bahauddin Zakariya University, Pakistan

Abstract

Presenting research is as important as conducting research; requiring novice researchers to practice and review their presentations continuously. Given the limited time in language classrooms, developing impactful presentation skills is difficult. Tools like AI chatbots could help overcome this challenge. While research shows AI tools aid in developing oral proficiency, there is limited empirical evidence on the effectiveness of voice-enabled chatbots for improving academic presentation skills. This study investigates the effectiveness of a voice-enabled chatbot 'Oratio', in enhancing academic presentation skills, mainly the lexicogrammatical and structural features. The study also explores the possibility of Oratio reducing the students' anxiety related to giving academic oral presentations. The participants of this experimental study were 20 Master of Philosophy, first and second year students from the English department at a public university in Pakistan. Pre-intervention and post-intervention presentations were assessed using a rubric designed by the researchers. The experimental group used Oratio and the control group did not use any interventions to support their presentation skills improvement. The Wilcoxon signed rank test showed no significant improvement in the pre-test and post-test results in the control and experimental groups; however, there is a marked improvement reflected by the median scores in the post-intervention performance for all the participants. Using Oratio appears to have helped students build their confidence by improving lexicogrammatical

E-mail: aisha.aslam@uq.edu.au & saqib@bzu.edu.pk

and structural issues in their presentations. Although this pilot study has limitations, it illustrates the potential of using voice-enabled AI chatbots in improving academic oral presentation skills of novice researchers.

Keywords: oral presentation, academic speaking, AI chatbots

Introduction and Background

There is an increasing expectation from researchers to not only disseminate their research with other scientists but also with a wider, non-expert audience (Lamarche et al., 2024; Plo-Alastrué & Corona, 2023). Not being able to naturally communicate well should not stop researchers from learning the craft of delivering engaging presentations (Youknovsky & Bowers, 2024). It is crucial that university students are taught to share their research with a wider audience in an engaging and accessible manner. AI chatbots can play a vital role in providing the support that is needed (Belda-Medina & Calvo-Ferrer, 2022; Belda-Medina & Kokošková, 2023; Labadze et al., 2023; López-Minotta et al., 2025; Pituxcoosuvann et al., 2025).

Research on AI chatbots has shown their potential in developing English language learners' speaking skills. A systematic review conducted by Du and Daniel (2024) focused on the literature (published between 2017 and 2023) on English-speaking skills improvement using AI chatbots, and concluded that using AI chatbots has generally proved to be effective in vocabulary (Chen et al., 2023) and grammar learning (Tai, 2024); confidence building (Zou et al., 2023); alleviating anxiety (Han, 2020; Tai & Chen, 2023); speaking fluency (Wu et al., 2023) and improving learners' motivation. There are a few studies done to date, investigating the improvement of oral presentation skills using AI chatbots. Garcia et al. (2024) studied the effectiveness of AI chatbots on reducing public speaking anxiety and improving self-perceived confidence and reported positive results. Another small-scale, preliminary study attempted to use AI chatbots along with the physiological detection model to help reduce anxiety in students while giving presentations (Kimani & Bickmore, 2019).

However, the research on the effectiveness of AI chatbots has shown certain limitations. For example, the studies have reported challenges in speech recognition leading to lack of accuracy of feedback given by the AI chatbots (Chen et al., 2023; Moussalli & Cardoso, 2016) resulting in lack of user satisfaction (Tai, 2024). The other limitation of AI chatbots is the occasional comprehensibility issues between the chatbots and the learners (Chen et al., 2023; Wu et al., 2020).

The lack of empirical evidence on using chatbots to improve lexicogrammatical and structural features in academic presentations of students in tertiary education has motivated this pilot study. Therefore, the researchers recruited the MPhil students to use a voice-enabled AI chatbot, Oratio, to help them practice their academic presentations. ‘Oratio’ is developed by the researchers with the help of a developer, and its efficacy is tested to answer the following questions:

- RQ1. How effective is Oratio in improving the research students’ lexicogrammatical features, structure, clarity, confidence, and engagement in academic presentations?
- RQ2. Does Oratio help students reduce their anxiety related to delivering academic presentations?

Methodology

The main objective of this study was to assess the effectiveness of a voice-enabled AI chatbot to improve students’ academic presentation skills, therefore, ‘Oratio’ – a voice-enabled chatbot, was developed because most voice-enabled chatbots are behind paywalls. The chatbot in its beta version is developed using Open AI’s GPT-4o-mini model. The current version takes audio input; therefore, the students can practice delivering their presentation in a safe, stress-free manner. Its feedback in this version is text-only, but voiced output can be an added feature in its full version. We decided to go against using the voiced feedback at this stage as the freely available chatbot voices are very robotic. Once funding is secured for this project, a full version of Oratio will have more realistic voiced as well as written feedback. Oratio gives feedback on vocabulary, grammar, structure, clarity, and engagement of the presentation based on the features of TED Talks.

The data were collected as oral presentations delivered by the MPhil English students in two groups: first year and second year students. Each group was divided into two sub-groups: (a). the Experimental group that used the voice-enabled chatbot ‘Oratio’, and (b). the Controlled group that did not receive any instructions or feedback. After the pre-intervention presentations, the latter group was suggested to watch some examples of good presentations on YouTube, such as TED Talks while the Experimental group used Oratio as their intervention. The students in first year presented a topic from a published research paper while the second year students presented their own research as they were already working on their MPhil thesis.

Table 1. Examination with pre- and post-tests across first and second year Mphil students

Groups	<i>N</i>	Median	<i>Z</i> value	<i>p</i> value	<i>r</i>
Pretest 1 st year control group	5	15	−1.76	0.7	−1.795
Post-test 1 st year control group	5	17			
Pre-test 1 st year experimental group	5	17	−1.84	.06	−1.718
Post-test Pretest 1 st year experimental group	5	21			
Pre-test 2 nd year control group	5	15	−.67	.50	−4.719
Post-test 2 nd year control group	5	18			
Pre-test 2 nd year experimental group	5	16	−1.84	.06	−1.718
Post-test 2 nd year experimental group	5	18			

Using the rubric designed by the researchers (see Appendix A), students' presentations were evaluated. The Wilcoxon signed-rank test was performed to analyze the differences in students' presentations, pre- and post-interventions.

Results and Discussion

The results of the Wilcoxon signed rank test did not show any significant improvement in the pre-test and post-test results in control and experimental groups; however, there is a marked improvement reflected by the median scores in the post-intervention performance across all the groups (see Table 1).

Structure of Presentations

Most of the students ($n = 7/10$) showed structural awareness in presentations. It is through the use of sentence structure that clearly indicates the part of the presentation being presented. The following examples from the first and second year experimental group show this. For example, they clearly

mentioned the sections in their presentation, as can be seen in the examples below:

(a) *Then there is an introduction.*

Then there is a difference of the EMI, ICLH and CLIC:

Then there is an English medium paradigm which is a ... (HA Pre-T-1)

The second part *is concerned with the rise of English education.*

Then the next part is *concerned with the key differences. (HA Post-T-1)*

Grammatical Accuracy/Complexity

More than half of the students ($n = 6/10$) used grammatically complex sentences in the experimental group. These students used complex sentence structures with appropriate use of connectors or linking words. In the example below, the first year student used a simple grammatical structure to describe ‘code switching’ in the pre-test, whereas in the post-test, the same participant used a complex grammatical structure to describe ‘code switching’.

(b) *Code switching is a very interesting [...]* (NA Pre-T-1)

So as far as the usage of code switching is concerned, *it is used [...]*

(NA Post-T-1)

Vocabulary Range and Appropriacy

The experimental group, second year, also showed variation in the use of vocabulary post-intervention.

(c) *Today, let's **delve into** the fascinating evolution of dictionaries. I'll **delineate** how dictionaries have evolved over time, specifically the shift from prescriptive to descriptive approaches in lexicography.* (AL Post-T-2)

Overall Clarity and Impact

Majority of the students in the experimental group ($n = 7/10$) showed improvement in overall clarity in presentations. The following extracts show that in the post-intervention, the student has conveyed the same ideas with more clarity.

- (d) *So to conclude that replication is vital yet underdeveloped in applied linguistics [...]* (MA Pre-T-1)
So replications hold a lot of significance for the applied linguistics [...] (MA Post-T-1)

Audience Engagement and Persuasion

More than half the students in the experimental groups ($n = 6/10$) displayed awareness in engaging the audience. The following example (e) shows that the experimental group student has used the phrase “*as you see*” to engage the audience.

- (e) *We see that the present study... We see the findings that the words like Janak, [...]*as well.** (AA Pre-T-1)

For example, we see in the findings that [...] frequency as you see in the article (AA Post-T-1)

Use of Metaphor, Simile, or Analogy

A few students in the experimental group ($n = 4/10$) made a use of analogies to explain the content of their presentation. In example (f) the student has used an analogy to make their presentation interesting.

- (f) *Myself, [name, roll number]. Today I am going to explore the topic, The English [...] Education.* (HS Pre-T-2)
My name is [...] Imagine navigating a vast ocean of knowledge and language is a vessel that carries us through. (HS Post-T-2)

The findings of the current study show that the students who used Oratio tried to make their presentations more interactive and interesting (see Appendix B for screenshots of students’ interaction with Oratio). The use of Oratio has helped in reducing their anxiety related to presentation. These results are consistent with the findings of the previous research (Askari & Rahim, 2024; Kimani & Bickmore, 2019). This could be explained in the light of Krashen’s (1982) Affective Filter Hypothesis and Sweller’s (2011) Cognitive Load Theory. More practice coupled with structured and personalized feedback may have reduced affective filter and lower the cognitive load for the students. The control group has shown little improvement in the use of vocabulary, structure, and engagement with the audience.

As their presentations lacked overall clarity and impact, and they did not try to engage with the audience, they were less confident and descriptive in their presentation. A possible reason could be the lack of awareness of the various features of academic presentation.

Conclusion

It is important that postgraduate students learn to improve their academic presentations. In the current study, the use of Oratio, a voice-enabled chatbot, has shown some improvement in the structure and confidence of the novice researchers. There are various limitations and challenges that this study faced. The small size of participants, discrepancy in tasks given to first year and second year students, the short time available to the Oratio group participants for practising their presentations before the post-tests, and dealing with natural disaster and unforeseen political tension with the neighbouring country were a few challenges that the researchers had to face during the data collection process. Above all, the AI chatbot, Oratio, was developed by the researchers on a self-finance basis. Due to limited funds, Oratio could not be designed as per the needs of the learners. We plan on using WhisperAI, deepgram and GPT-5 in the full version of Oratio to enable us to add better voiced output and playback along with the text feedback. We would also improve the UX design of Oratio and may develop a companion mobile app.

References

- Askari, M., & Rahim, A. S. (2024). The role of AI in enhancing presentation skills in English language teaching. *International Journal of Business and Applied Social Science*, 10(12), 1–6. <https://doi.org/10.33642/ijbass.v10n12p1>
- Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. *Applied Sciences*, 12(17), Article 8427. <https://doi.org/10.3390/app12178427>
- Belda-Medina, J., & Kokošková, V. (2023). Integrating chatbots in education: Insights from the Chatbot-Human Interaction Satisfaction Model (CHISM). *International Journal of Educational Technology in Higher Education*, 20(1), Article 62. <https://doi.org/10.1186/s41239-023-00432-3>
- Chen, H. H. J., Yang, C. T. Y., & Lai, K. K. W. (2023). Investigating college EFL learners' perceptions toward the use of Google Assistant for foreign language learning. *Interactive Learning Environments*, 31(3), 1335–1350. <https://doi.org/10.1080/10494820.2020.1833043>

- Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence*, 6, Article 100230. <https://doi.org/https://doi.org/10.1016/j.caeai.2024.100230>
- Garcia, F. V., Jr., Pascual, N. F., Sybingco, M. E., & Ong, E. (2024). Addressing public speaking anxiety with an AI speech coach. In A. Kashiwara et al. (Eds.), *Proceedings of the 32nd International Conference on Computers in Education*. Asia-Pacific Society for Computers in Education. <https://doi.org/10.58459/icce.2024.4955>
- Kimani, E., & Bickmore, T. (2019). Addressing public speaking anxiety in real-time using a virtual public speaking coach and physiological sensors. In *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents (IVA '19)* (pp. 260–263). Association for Computing Machinery. <https://doi.org/10.1145/3308532.3329409>
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon.
- Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education*, 20(1), Article 56. <https://doi.org/10.1186/s41239-023-00426-1>
- Lamarche, V. M., Kung, F. Y. H., Finkel, E. J., Jayawickreme, E., Rattan, A., & Wheatley, T. (2024) How to give great research talks to any audience. *Nature Human Behaviour*, 8, 616–621. <https://doi.org/10.1038/s41562-024-01839-2>
- López-Minotta, K. L., Chiappe, A., & Mella-Norambuena, J. (2025). Implementation of Artificial Intelligence to improve English oral expression. *Multidisciplinary Journal of Educational Research*, 15(1), 43–71. <https://doi.org/10.17583/remie.16188>
- Moussalli, S., & Cardoso, W. (2016). Are commercial ‘personal robots’ ready for language learning? Focus on second language speech. *CALL Communities and Culture—Short Papers from EUROCALL*, 325–329. <https://doi.org/10.14705/rpnet.2016.eurocall2016.583>
- Pituxcoosuvann, M., Tanimura, M., Murakami, Y., & White, J. S. (2025). Enhancing EFL speaking skills with AI-powered word guessing: A comparison of human and AI partners. *Information*, 16(6), Article 427. <https://doi.org/10.3390/info16060427>
- Plo-Alastrué, R., & Corona, I. (Eds.). (2023). *Digital scientific communication: Identity and visibility in research dissemination*. Springer.
- Sweller, J. (2011). Cognitive load theory. In *Psychology of learning and motivation* (Vol. 55, pp. 37–76). Academic Press.

- Tai, T. Y. (2024). Effects of intelligent personal assistants on EFL learners' oral proficiency outside the classroom. *Computer Assisted Language Learning*, 37(5–6), 1281–1310. <https://doi.org/10.1080/09588221.2022.2075013>
- Tai, T. Y., & Chen, H. H. J. (2023). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments*, 31(3), 1485–1502. <https://doi.org/10.1080/10494820.2020.1841801>
- Wu, C., Lam, H., Kong, A., & Wong, W. (2023). The design and evaluation of a digital learning-based English chatbot as an online self-learning method. *International Journal of Engineering Business Management*, 15. <https://doi.org/10.1177/18479790231176372>
- Wu, E. H. K., Lin, C. H., Ou, Y. Y., Liu, C. Z., Wang, W. K., & Chao, C. Y. (2020). Advantages and constraints of a hybrid model K-12 E-Learning assistant chatbot. *Ieee Access*, 8, 77788–77801. <https://doi.org/10.1109/ACCESS.2020.2988252>
- Youknovsky, A., & Bowers, J. (2024). The SELL method: Prepare your presentation. In A. Youknovsky & J. Bowers (Eds.), *SELL YOUR RESEARCH: Public Speaking for Scientists* (pp. 5–14). Springer International Publishing. https://doi.org/10.1007/978-3-031-60887-2_2
- Zou, B., Du, Y., Wang, Z., Chen, J., & Zhang, W. (2023). An investigation into Artificial Intelligence speech evaluation programs with automatic feedback for developing EFL learners' speaking skills. *SAGE Open*, 13(3). <https://doi.org/10.1177/2158244023119381>

Appendix A

The Rubric to Evaluate Presentations

Criterion	Level 1 (Limited)	Level 2 (Developing)	Level 3 (Proficient)	Level 4 (Strong)
Structure of Presentation	<i>Lacks clear beginning, body, and end. Ideas are disjointed.</i>	<i>Attempt at structure but transitions are unclear; some logical flow.</i>	<i>Clear structure with recognizable introduction, body, and conclusion. Transitions mostly smooth.</i>	<i>Strong and cohesive structure with well-signposted transitions; smooth flow throughout.</i>
Grammar Accuracy (CEFR A2–B2)	<i>Frequent errors impede understanding (A1–A2 level).</i>	<i>Some basic grammatical control, though errors are noticeable (low B1).</i>	<i>Generally good control of simple and some complex forms; few errors (mid–high B1).</i>	<i>Uses a range of grammatical structures with high accuracy and flexibility (B2+).</i>
Vocabulary Range and Appropriacy (CEFR A2–B2)	<i>Limited vocabulary; repetitive or incorrect word choices; lacks academic register.</i>	<i>Some variety; attempts to use academic vocabulary; occasional misuse.</i>	<i>Appropriate range of vocabulary, including some academic terms; mostly accurate.</i>	<i>Wide lexical range; precise and appropriate use of academic and topic-specific terms.</i>
Confidence in Delivery	<i>Hesitant, reads from notes/script; avoids eye contact; posture closed.</i>	<i>Some hesitations, occasional eye contact; limited engagement with audience.</i>	<i>Generally confident, maintains eye contact; uses gestures; some natural delivery.</i>	<i>Speaks fluently with clear articulation; confident body language and tone; highly engaging.</i>
Overall Clarity and Impact	<i>Difficult to follow; lacks coherence or persuasiveness.</i>	<i>Somewhat clear but lacks impact or persuasiveness.</i>	<i>Clear and coherent; makes an effort to engage the audience.</i>	<i>Clear, impactful, and engaging; demonstrates persuasive communication.</i>
Audience Engagement and Persuasion	<i>Monotonous or disengaged; lacks connection with the audience.</i>	<i>Occasional attempts to engage the audience.</i>	<i>Generally engaging; uses some rhetorical questions or hooks to persuade or connect.</i>	<i>Highly engaging and persuasive; uses rhetorical questions, 'survey questions' and/or hooks effectively.</i>
Use of Metaphor, Simile, or Analogy	<i>No figurative language used; presentation is literal and dry.</i>	<i>Rare use of metaphor/simile; not clearly connected to content.</i>	<i>Effective use of metaphor or analogy in parts of the presentation.</i>	<i>Creative and well-integrated use of metaphor/simile/analogy that enhances understanding and engagement.</i>

Note. The inter-rater reliability of the rubric is tested for both pre-test and post-test evaluations. Pre-test scores was examined using a two-way mixed-effects model of the intraclass correlation coefficient (ICC), assessing consistency between two evaluators' ratings of speaking proficiency. The analysis of pre-test evaluation showed a single-measures ICC of .85, 95% CI [.67, .94], indicating substantial agreement when considering one rater alone. The average-measures ICC was .92, 95% CI [.80, .96], reflecting excellent reliability. Similarly, the analysis of post-test evaluation showed a single-measure ICC of .93, 95% CI [.84, .97], indicating excellent agreement when considering one rater alone. The average-measures ICC was .96, 95% CI [.91, .98], reflecting excellent reliability.

Appendix B

The screenshots from the chatbot- students interaction may be accessed here.

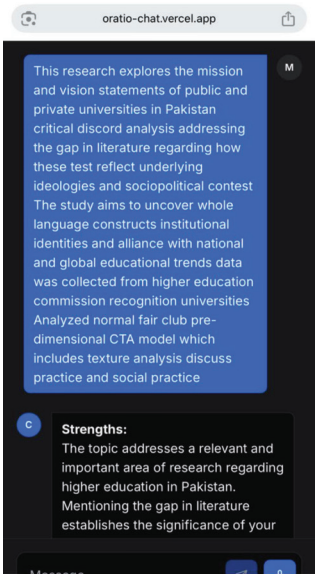


Figure 1(a). Student A's input high-lighted in blue

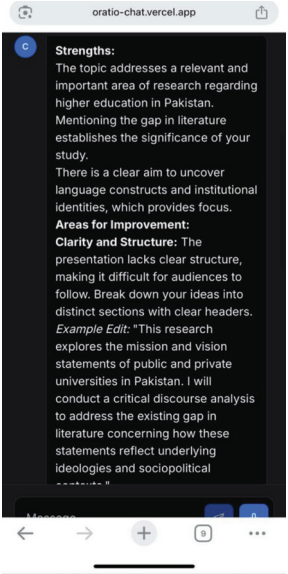


Figure 2. Oratio's feedback on student A's input

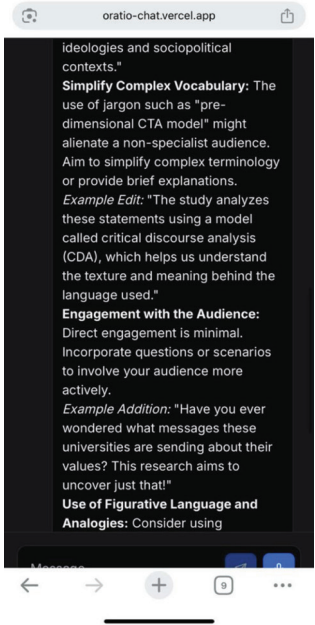


Figure 3. Oratio's feedback on student A's input (cont.)

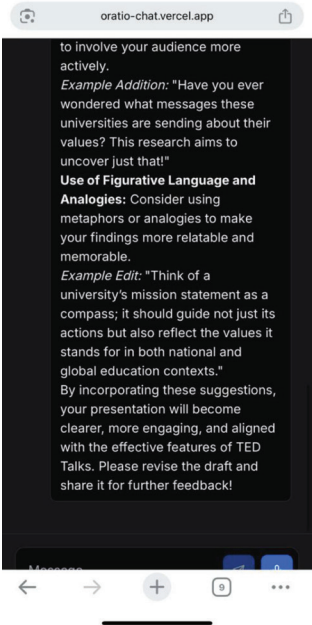


Figure 4. Oratio's feedback on student A's input (cont.)

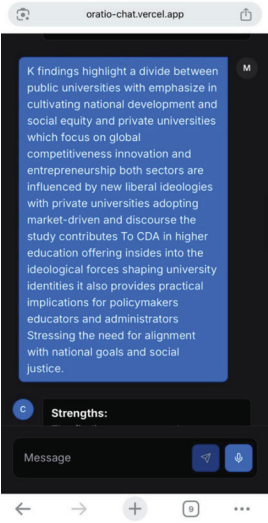


Figure 5(a). Student B's input highlighted in blue

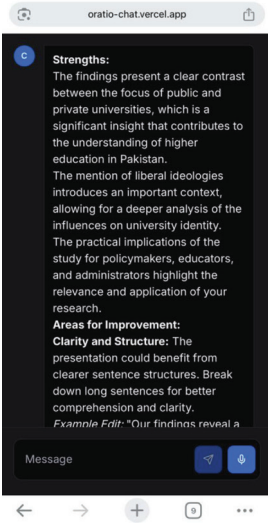


Figure 6. Oratio's feedback on student B's input

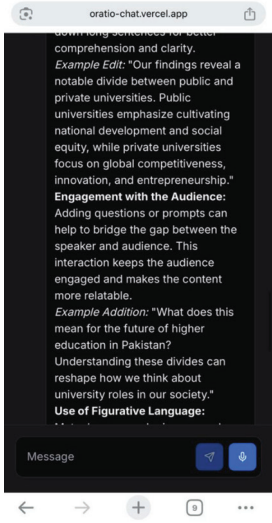


Figure 7. Oratio's feedback on student B's input (cont.)

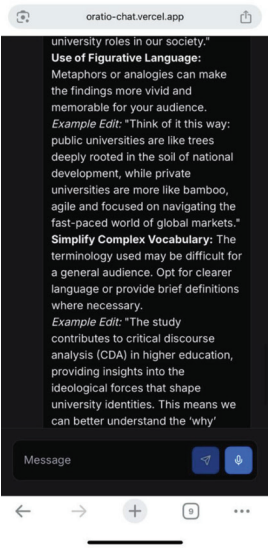


Figure 8. Oratio's feedback on student B's input (cont.)

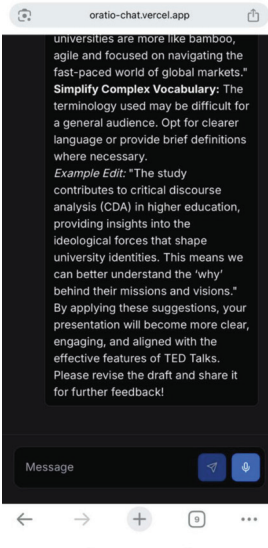


Figure 9. Oratio's feedback on student B's input (cont.)