African Journal of Advanced Studies in Humanities and Social Sciences (AJASHSS)
المجلة الإفريقية للدراسات المتقدمة في العلوم الإنسانية والاجتماعية
Online-ISSN: 2957-5907
Volume 5, Issue 2, 2026
Page No: 673-678
Website: https://aaasjournals.com/index.php/ajashss/index
Arab Impact factor 2025: 0.83 **SJIFactor 2024: 6.856** **ISI 2025: 1.191**

AI-Driven Interpreting Training: Students' Perceptions of ChatGPT as a Spoken Trainer with Human Guidance

Mohamed Ali Suliman Abbas *
Translation Department, Faculty of Languages and Translation, Misurata University, Misurata, Libya

التدريب على الترجمة الشفهية باستخدام الذكاء الاصطناعي: آراء الطلاب في تشات جي بي تي بلس مدربا صوتيا بإشراف بشري

محمد علي سليمان عباس *
قسم الترجمة، كلية اللغات والترجمة، جامعة مصراتة، مصراتة، ليبيا

*Corresponding author: m.abbas@lt.misuratau.edu.ly

Received: March 28, 2026	Accepted: June 04, 2026	Published: June 20, 2026

Abstract:
All disciplines of scientific enquiry have fruitfully deployed artificial intelligence, and language education is no exception. In this study, twenty participants from the Translation Department at Misurata University, including both recent graduates and advanced students who had successfully completed their core interpreting coursework, explored their perception of using ChatGPT Plus as a spoken interpreting trainer within a structured, human-guided framework. The research employed a mixed-methods approach utilizing a comprehensive survey designed to capture both quantitative trends and qualitative perceptions. This instrument assessed the tool's effectiveness in supporting vocabulary acquisition, improving reflexes, translation accuracy, and both consecutive and simultaneous interpreting performance. Findings revealed that students perceived the tool as helpful and engaging, particularly for consecutive interpreting training, as ChatGPT Plus provided participants with flexible, self-paced oral practice. However, technical issues hindered simultaneous interpreting training. The tool always stopped delivering speech when participants began interpreting, mistaking their rendition for a conversational command. Other limitations included early task termination and overly positive automated feedback that occasionally overlooked translation inaccuracies. The findings highlight that while AI tools can improve student motivation and support mechanical skill development, they are most effective when integrated into a framework of consistent human guidance.

Keywords: Artificial Intelligence, ChatGPT Plus, Interpreter Training.

الملخص
لقد وظفت شتى ميادين البحث العلمي الذكاء الاصطناعي توظيفاً مثمراً، وليس تعليم اللغات بمنأى عن هذا التوجه. تناولت هذه الدراسة عشرين مشاركًا من قسم الترجمة بجامعة مصراتة، بينهم خريجون حديثون وطلبة متقدمون أنهوا مقررات الترجمة الشفهية. استهدفت الدراسة استكشاف تصوراتهم في استخدام تشات جي بي تي بلس مدرّبا صوتيا للترجمة الشفهية ضمن إطار منظم وبإشراف بشري. اعتمد البحث منهجًا مختلطًا، مستخدمًا استبيانًا مفصلًا لجمع البيانات الكمية والآراء النوعية. وقد قيّم الاستبيان دور الأداة في بناء المفردات، وتطوير سرعة الاستجابة، وتحسين دقة الترجمة، وتعزيز الأداء في الترجمة التتابعية والفورية. خلصت النتائج إلى أن الطلاب وجدوا الأداة نافعة وتفاعلية، خاصة في التدريب على الترجمة التتابعية، حيث وفّرت لهم ممارسة شفهية مرنة وذاتية. لكن المشكلات التقنية حَدَّت من فعاليتها في الترجمة الفورية؛ إذ كان النظام يتوقف عن الكلام عند بدء الطلاب بالترجمة الفورية، ظَنًا منه أن كلامهم (ترجمتهم) عبارة عن أوامر صوتية. كما ظهرت عيوب أخرى مثل إنهاء التدريب مبكراً، والإفراط في المديح وإغفال الأخطاء أحيانا. وتوضح النتائج أن أداة تشات جي بي تي بلس تزيد من حماس الطلاب وتدعم تطوير مهاراتهم الأساسية، لكنها ستكون أكثر فاعلية عندما يصحبها توجيه وإشراف بشري مستمر.

الكلمات المفتاحية: الذكاء الاصطناعي، تشات جي بي تي بلس، تدريب المترجمين الشفهيين.

Introduction

Artificial Intelligence (AI) has come to stay. It is a permanent reality and no longer a futuristic prospect. (Baidoo-Anu & Ansah, 2023). Given this reality, any field that does not make an alliance with it is bound to fall behind. In language teaching, AI is being put into practical uses, and so is now the case in the interpreting training domain (Sánchez-Castany, 2025; Van Horn, 2024). AI applications have been used in conversational chatbots, automated writing evaluation systems, and more. Such applications demonstrated both promise and limitations (Handley, 2024; Van Horn, 2024). However, the specific application of large language models to interpreting training remains relatively under-researched. That may be attributed to the fact that interpreting demands real-time processing, rapid decision-making, and the ability to manage cognitive load while managing two languages simultaneously. This creates a unique pedagogical context.

Large language models (LLM) like ChatGPT offer promising possibilities for this specialized training. They can generate realistic spoken discourse, and provide immediate feedback. They also adapt to individual learning preferences, and offer unlimited practice opportunities even when there is no human trainer around. These automated capabilities, however, require investigation to find out if they translate into effective learning outcomes in a live, spoken environment.

Research Objectives

1- To evaluate student perceptions regarding the effectiveness of ChatGPT Plus as an oral interpreting trainer.
2- To identify the pedagogical strengths and technical limitations of using AI in consecutive and simultaneous interpreting practice.

Research Questions

1- How do translation students perceive the effectiveness of ChatGPT Plus in supporting vocabulary acquisition and interpreting reflexes?
2- How do translation students perceive and experience the technical challenges of AI-driven (ChatGPT Plus) simultaneous interpreting compared to consecutive interpreting?

Literature Review

The Evolution of Automated Feedback in Interpreting

Early exploration into computer-assisted interpreter training mainly focused on static automation. In other words, they relied on speech recognition and machine translation for error analysis and vocabulary tracking. Scholars like Fantinuoli (2018) and Hatiarová (2025) showed that these early systems were useful for building specialized glossaries and identifying lexical errors. On the negative side, such systems struggled to process cultural nuances or deliver fluid feedback. Lin (2024) later noted that technology could help simulate interpreting environments without needing human actors. But these pre-LLM tools lacked any real capacity for interactive or personalized dialogue. Thus, early software remained a supplementary tool for mechanical practice rather than an active training partner.

The Transition to Interactive Large Language Models

The shift towards LLMs basically changed this dynamic by introducing platforms capable of responsive, conversational dialogue. In the context of language education, early LLM research focused heavily on automated text feedback and customized lesson planning (Baidoo-Anu & Ansah, 2023). For instance, Hatiarová (2025) demonstrated that ChatGPT could generate tailored, domain-specific interpreting scenarios based on student needs. With such capabilities taken into account, researchers quickly identified that these conversational breakthroughs need human supervision. They cautioned that LLMs frequently generated subtle inaccuracies and lacked the pedagogical grounding required for independent professional training (Hatiarová, 2025).

Human-Guided AI and Interpreter Pedagogy

Current research focuses on blending these generative AI tools with structured human instruction. This way, learning environments are thought to be more balanced. Recent studies show that students highly value the constant accessibility of AI platforms, particularly for self-paced, consecutive interpreting practice (Lin, 2024). Even so, contemporary pedagogical frameworks emphasize that AI should supplement rather than replace a lecturer's expertise. Cui et al. (2025) underscore that students still depend entirely on human instructors to receive nuanced feedback on cultural sensitivity and complex simultaneous interpreting techniques.

Methodology

Research Design

To truly understand how an AI-driven environment reshapes the interpreting classroom, this study moves beyond simple numerical data. The research adopts a mixed-methods design, blending quantitative results with qualitative

student reflections to provide a more complete picture of the learning process. This study intentionally adopts a user-centered perception framework. It does not follow a traditional pre- and post-test linguistic assessment. Understanding students' attitudes, psychological safety, direct interface interactions is a crucial prerequisite before any empirical testing can be meaningly designed. This integrated framework ensures that the findings are statistically sound. It also guarantees that the data remains deeply reflective of the real-world operational challenges students face when practicing with ChatGPT Plus.

Participants

The sample for this study comprised 20 participants from the Translation Department at Misurata University. This group included advanced undergraduate students in their final semester and recent graduates who had recently completed the program. To ensure a high level of competency, all participants were required to have completed the interpreting courses required by the programme. The study was conducted during the Autumn 2025–2026 academic term.

3.2 Procedure

The research was carried out over a three-week period. During this time, participants engaged in structured training sessions using ChatGPT Plus as a spoken interpreting trainer. The curriculum was specifically designed to bridge the gap between theory and practice, targeting vocabulary acquisition, translation accuracy, and the development of interpreting reflexes.

The training protocol followed a scaffolded, three-stage workflow:

1. Lecturer-Provided Materials: To ground the AI in professional standards, the instructor first fed ChatGPT Plus with specific materials, including targeted glossaries, vocabulary lists, and thematic texts. Detailed prompts were used to ensure the ChatGPT Plus followed a structured pedagogical path.

2. Scaffolded Practice Sessions: Students progressed through a hierarchical learning loop:

    - Vocabulary Phase: Participants were first presented with target terminology and their translations, followed by a quick testing phase to ensure retention.
    - Sentence Phase: Students practiced interpreting individual sentences. Crucially, ChatGPT Plus was instructed to slightly modify the syntax of these sentences in real-time while maintaining the core vocabulary, forcing students to adapt their reflexes.
    - Spoken Passage Phase: In the final stage, students interpreted short, spoken speeches based on the previously taught topics.

3. Human-Guided Mediation: The instructor acted as a technical and educational facilitator, monitoring the digital interaction to maintain the session's integrity. This involved intervening when ChatGPT Plus strayed from the instructions, managing the AI's tendency to stop the dialogue too early, or resolving any voice-recognition errors that disrupted the students.

Data Collection

Primary data collection relied on post-session surveys distributed immediately after each one-hour training block. Students completed a comprehensive questionnaire utilizing Likert-scale items to measure the tool's effectiveness and practicality. To ensure a complete picture, the survey also included open-ended prompts, allowing participants to describe their direct engagement and share technical frustrations in their own words.

Data Analysis

The analysis followed a dual-track strategy to respect the mixed-methods nature of the study:

- Quantitative Analysis**:** Descriptive statistics were applied to the Likert-scale data to establish general trends in student perceptions regarding the tool's effectiveness across different interpreting modalities.

- Qualitative Analysis**:** The qualitative data, gathered from the open-ended sections of the post-session surveys, were analyzed using thematic analysis. This process involved identifying recurring patterns in student feedback. Following the pioneering six-phase framework established by Braun and Clarke (2006), this process involved a systematic coding approach to identify, analyze, and report recurring patterns and core experiences shared by the participants.

Results

This study evaluated how effectively ChatGPT Plus functions as a spoken trainer for interpreting students. The findings, drawn from both the quantitative survey responses and the students' open-ended reflections, show a clear divide in how the tool performed depending on the mode of interpreting practiced.

Quantitative Performance Metrics

As shown in Table 1, participants found ChatGPT Plus to be highly effective for vocabulary preparation and consecutive interpreting tasks. However, its ratings dropped sharply when applied to simultaneous interpreting.

Table 1: Student Perception of Skill Improvement and Tool Effectiveness (N=20)

Domain of Competency	Mean Score (1–5)	SD	Perception Category
Vocabulary Acquisition	4.82	0.35	Exceptional
Consecutive Rendition	4.60	0.48	Highly Effective
Interpreting Reflexes	4.15	0.61	Effective
Simultaneous Interpreting	2.15	0.92	Inadequate
Overall Engagement	4.55	0.42	High

Qualitative Synthesis

The thematic analysis of the open-ended survey questions revealed three central experiences shared by the participants:

1. A Low-Anxiety Practice Space: A striking 90% of the students highlighted that practicing with ChatGPT Plus felt incredibly safe. Because they were interacting with a machine rather than a professor, the normal fear of making mistakes in front of others largely disappeared, allowing them to practice more freely. These qualitative insights justify the choice of a perception-based study. An empirical test might show if a student's vocabulary improved, but it would fail to capture the critical psychological shift. An example for this is the 90% reduction in performance anxiety reported by participants, which is essential to independent interpreter training.
2. The Voice Interruption Flaw: During simultaneous interpreting tasks, students encountered a frustrating and recurring technical error. ChatGPT Plus frequently mistook the student's voice for a conversational command. Thus, it kept stopping the speech and breaking the flow of the exercise.
3. Fluency vs. Accuracy Errors: Students noticed that while ChatGPT Plus gave feedback almost instantly, its praise could be misleading. It often complimented translations that sounded smooth and fluent, even when those translations contained clear errors in terminology or meaning.

Discussion

The core finding from this study is that ChatGPT Plus is not a one-size-fits-all tool for interpreting students. Its value depends entirely on what skill is being practiced. The findings show that the app performs much better as a consecutive interpreting trainer than a simultaneous one. Nonetheless, consecutive interpreting and vocabulary building are not isolated activities; they are the very stepping stones students need to become good simultaneous interpreters. By excelling at consecutive practice, ChatGPT Plus essentially prepares students for the heavier cognitive demands of simultaneous work, even if the software itself hits a technical wall later on.

Building the Foundation: Vocabulary and Consecutive Practice

The survey data showed an incredibly high satisfaction score for vocabulary acquisition (M=4.82) and consecutive training (M=4.60). During the sessions, students felt that ChatGPT Plus acted like a patient, private partner. They did not have to worry about a professor evaluating them or classmates judging their hesitations. This psychological safety let them focus all their mental energy on note-taking, listening, and retrieving words. One advanced student explained this beautifully: "With the app, [the] pressure was gone. I practiced the same difficult passage four times until my tongue got used to [the] term[s]." Because mastering consecutive delivery and automating vocabulary are the direct prerequisites for simultaneous interpreting, this private drilling is invaluable. It gives students the exact baseline confidence and rapid reflexes (M=4.15) they need before the training gets complicated.

The Structural Divide: Evaluating Short Units vs Long Continuous Speech

A notable contradiction emerges when contrasting the high satisfaction scores for vocabulary acquisition (M=4.82) and interpreting reflexes (M=4.15) against reported "politeness bias" where the software sometimes overlooked some translation errors. This paradox is resolved by examining the size of the linguistic unit under practice. ChatGPT Plus proved highly capable and reliable when assessing individual words in isolated terminology during the initial vocabulary drills. At this micro-level, its automated feedback was more precise. The accuracy dropped, however, when students progressed to interpreting full sentences and particularly in long

consecutive tasks. In those longer modalities, the software seemed to prioritize fluency and smooth delivery over deep semantic and syntactic accuracy. Therefore, the tool's assessment is more precise when dealing with smaller chunks of language and less accurate with longer ones.

The Technical Wall: Moving into Simultaneous Interpreting

The problem arose when students tried to take this solid vocabulary foundation and apply it to live simultaneous practice, where satisfaction scores dropped drastically to an inadequate level (M=2.15). This drop did not happen because the students lacked the skills. This happened because the technology physically blocked them. ChatGPT Plus is engineered for one-way alternating dialogue. In other words, one person talks, stops, and then the other responds. Simultaneous interpreting requires the exact opposite. The student must speak and listen at the same time. The moment a student began interpreting over the live source audio, ChatGPT Plus voice recognition mistook the student's voice for an interruption. It assumed the user was changing the prompt and went completely silent. This structural barrier made it impossible for students to practice. A recent graduate detailed how frustrating this was: "As soon as I began interpreting, the AI started responding. It thought I was giving instructions." So, while the tool is incredibly effective at building the underlying linguistic building blocks, the actual voice interface cannot handle the physical reality of simultaneous overlapping audio.

Prompt Defiance and Student Frustration

Despite high overall engagement scores (M=4.55), a notable sense of annoyance was reported by participants when the ChatGPT Plus refused to follow instructions. In interpreter training, pacing and duration are vital, such as demanding that the AI maintain a source speech for a full three minutes to test stamina. However, during live voice sessions, ChatGPT Plus frequently suffered from instruction drift, ignoring previous commands and ending exercises early. This forced students to interrupt their performance to reset the exercise parameters, leading to intense frustration. One participant noted the irritation of dealing with this stubbornness: "I felt angry because ChatGPT did not follow [my] orders [instructions]... I told it five times to give me 20 sentences to translate, but it stopped after four or five sentences."

The Danger of the "Politeness Bias"

Even within the successful vocabulary and consecutive drills, another limitation was discovered. ChatGPT Plus was often too polite to be a standalone trainer. Students noticed that the app consistently praised their delivery as long as it sounded smooth and fluent, completely missing severe translation errors. In professional translation, an error made with confidence is far more dangerous than a hesitant self-correction, because it completely misleads the audience. One participant shared a striking example of this: "Sometimes, I know my translation was not good, but I still got praise from the programme." This confirms that human subject-matter experts are needed to act as mediators—not just to double-check the AI's vocabulary accuracy, but to actively manage the software's behavior when it ignores pedagogical constraints.

Implications for Pedagogy

The practical lesson from this study is that we shouldn't view AI as a threat to the traditional interpreting lab, but rather as an excellent preparatory tool students can use before they ever step inside one. Instead of replacing the lecturer, the technology is best used to handle the heavy lifting of baseline drill work. To make this work practically in a university setting, we propose a simple, two-tiered approach:

- The Independent Practice Phase (AI-Driven): Students work at home with the LLM to handle high-repetition, mechanical tasks. This is where they build specialized glossaries, run quick syntactic reflex drills, and practice their self-paced consecutive renditions.
- The Mediated Classroom Phase (Human-Led): Valuable class time with the lecturer is saved for advanced simultaneous techniques and what we call the "art" of interpreting. This includes navigating complex cultural nuances, managing speaker register, and ensuring high-level accuracy. These are critical areas where the AI still falls short.

Conclusion

This research shows that ChatGPT Plus is a valuable asset for foundational interpreter training, especially when it comes to building vocabulary and mastering consecutive practice. It gives students a level of personalized, self-paced spoken interaction that was previously impossible without a one-on-one tutor. This way, independent practice has become much more accessible. However, the software's inherent design that follows a one-way, alternating conversation means it simply cannot serve as an independent trainer for simultaneous interpreting right now.

Limitations and Future Research

There are a few clear limitations to keep in mind when looking at these results. First, the study relies on a relatively small group of participants (N=20) working within a single linguistic context (Arabic-English) at Misurata University. Second, because this data looks at how the late-2025 version of ChatGPT Plus behaved, these findings capture a specific moment in a very fast-moving technological landscape.

Moving forward, future studies should look past student perceptions alone and use objective pre- and post-tests to measure actual skill improvement. Additionally, researchers should test newer conversational audio engines as they come out to see if future voice modes can finally fix the frustrating voice interruption errors that currently disrupt simultaneous practice.

References

[1] Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, *7*(1), 52–62. https://doi.org/10.61969/jai.1337500

[2] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

[3] Cui, F., Li, D., & Zhuang, C. (2025). Introduction: Transforming translation education through artificial intelligence. *The Interpreter and Translator Trainer*, *19*(3–4), 227–233. https://doi.org/10.1080/1750399X.2025.2561258

[4] Fantinuoli, C. (2018). Interpreting and technology: The upcoming of artificial intelligence. In C. Fantinuoli (Ed.), *Interpreting and technology* (pp. 1–12). Language Science Press.

[5] Handley, Z. (2024). Has artificial intelligence rendered language teaching obsolete? *The Modern Language Journal*, *108*(3), 548–555.

[6] Hatiarová, P. (2025). AI in interpreting training. *L10N Journal*, *1*(4), 45–66.

[7] Lin, Y. (2024). A study on AI-assisted and skill-based class training designed for English-Chinese consecutive interpretation. *Proceedings of the 3rd International Conference on Artificial Intelligence and Education (ICAIE 2024)*, 312–319. Association for Computing Machinery.

[8] Sánchez-Castany, R. (2025). The challenges of teaching translation technologies in the AI era. *Cadernos de Tradução*, *45*, Article e105121. https://doi.org/10.5007/2175-7968.2025.e105121

[9] Van Horn, K. R. (2024). ChatGPT in English language learning: Exploring perceptions and promoting autonomy in a university EFL context. *Teaching English as a Second Language Electronic Journal (TESL-EJ)*, *28*(1), Article 9. https://doi.org/10.55593/ej.28109a8