

Using AI Transcription-Translation Tools to Improve Speaking Skills for the IELTS Test

doi: 10.52987/edc.2025.006

Cristina Tat

International Studies Department, Doshisha Women's College of Liberal Arts

Abstract

This paper aims to share classroom practices for using the transcription-translation platform TurboScribe in conjunction with the latest Chat GPT-4o AI model to help students practice for the speaking component of the IELTS test. The participants are first-year undergraduate students in the International Studies Department at a private university in Western Japan. One year of study abroad is mandatory for these students, so they need to take the IELTS test to qualify for programs in North America, Europe, Southeast Asia, and Australia. The students' transcripts have been analyzed to check how they incorporated the advice received from the AI into their speaking practice. Students' perceptions and thoughts about using the AI, as well as the successes and challenges encountered during classroom instruction, are discussed.

Key Words: Artificial Intelligence (AI), ChatGPT, L2 speaking development, translation-transcription, feedback, IELTS test

Introduction

I. Program Overview

The mission of the International Studies department where this case study took place is to send all enrolled students for one year of study abroad. All the overseas affiliated institutions accepting these students use the IELTS test as a benchmark for acceptance. The cut-off score for acceptance is a band score of 5.5. As such, preparatory courses for the four skills tested on the IELTS, Listening, Speaking, Reading, and Writing, are required courses for first-year students. The speaking component of the IELTS test is taught in a year-long dedicated Speaking & Writing course with a maximum enrollment of fourteen students per class. The motivation for this study was to propose a guide for best teaching practices that would help these students to improve their IELTS speaking scores necessary for study abroad.

II. The Role of Transcription-Translation Tools in Fostering Speaking Practice

In spite of many years of English as a Foreign Language (EFL) study, many students struggle with production in L2 (Akhter et al., 2021; Kargar Behbahani et al., 2024, Leeming et al., 2024). For Japanese learners in particular, the heavy focus on grammar studies (Sakui, 2004) and the general tendency to sit quietly in lecture style, non-interactive courses, have led to a lack of active participation and engagement in the EFL classroom (Harumi, 2011; Kim, 2024). Therefore, it is important to offer tools to EFL students that will help them to visualize their speaking output, to train them to recognize errors and self-correct, and to provide feedback that is easy for them to understand and incorporate in future output. AI-powered tools that can help achieve these aims are now widely available to both students and educators, but studies on how CALL affects speaking efficacy are just beginning to emerge (Dong et al., 2022), as the use of AI technology is becoming more and more popular in language teaching (Dennis, 2024).

Transcription tools have been considered a great aid to help students improve both accuracy and fluency which are among the scoring criteria for the speaking component of the IELTS test. The IELTS speaking test consists of three parts: Part I is a basic interview about personal preferences and opinions on general topics; Part II is a one to two minutes monologue about a particular suggested topic with one minute of preparation; Part III requires generalized answers to more abstract themes based on the topic from Part II.

Research has shown that students can make significant gains in speaking ability by learning to spot and correct errors in transcribed speech in the short-term (Mennim, 2003; Stillwell et al., 2010) as well as in the long-term (Mennim, 2007). According to Stones (2013), transcription task cycles “[involving] practice test, transcription, student correction, teacher correction, and retrieval of the original test...showed that students’ approaches to correction varied widely from single corrections to complete reformulations, and that students were able to make gains with a variety of forms, appearing more likely to use their own corrections than those of the teacher.” Lynch (2007) has also shown that students’ perceptions of the process of using transcribed speech to analyze their speaking have been generally positive. A study carried out on Japanese learners of English by Hirai and Kovalyova (2024) has shown that speech-to-text applications demonstrate great pedagogical potential.

Methodology

I. AI Translation-Transcription Tools

The transcription platform selected for the IELTS Speaking & Writing course is TurboScribe, a Whisper powered AI that offers video and audio translation and transcription. The site allows three free transcriptions of up to 90 minutes per day and unlimited transcription-translation for paying subscribers. The resulting transcriptions can be downloaded and shared as PDF, DOCX, captions and subtitles (SRT/VTT), CSV, and TXT. Chat GPT-4o AI is linked to TurboScribe, so the transcripts can be directly analyzed through the same site. Instructors purchased paid accounts, while students only registered for free accounts.

II. Training for Students

To help students become accustomed to utilizing the TurboScribe and ChatGPT AI, training was provided. Students were shown how to make recordings on their devices and how to upload them to TurboScribe. Then they were taught how to download the transcribed text to Word documents which were printed. Instructor-created error-correction guides were introduced to explain how to spot common mistakes within the transcribed text and how to self-correct. The initial guides were simple, but they became more complex as students learned new test-taking techniques, formulaic language, and vocabulary. The instructor demonstrated how to use the error-correction guide through pair-work activities.

III. Transcription-Task Cycles

Speaking practice for the IELTS test was conducted in the following manner: after a new IELTS topic and techniques were introduced, students were given a set of Part I, II, and III questions on that topic for practice. Students were instructed to record their answers to the questions and transcribe the recorded speech using TurboScribe. They were provided with error-correction guides and asked to analyze and self-correct their speech. These assignments were carried out in class or assigned as homework. The self-corrected transcripts were then submitted to the instructor for feedback.

Before receiving instructor feedback, students were asked to feed their corrected transcription to ChatGPT and ask advice from the AI. The following suggested prompts were used for this purpose:

“How can I improve this answer for Part (I, II, or III) of the IELTS Speaking Test?”
“How can I improve this answer for Part (I, II, or III) of the IELTS Speaking Test in order to receive a 6.0 score?”

The second prompt was used after students had taken the IELTS for the first time and had an idea of their band score. Students would then compare their self-corrected transcripts with the suggestions offered by ChatGPT. Finally, the students received the self-corrected transcripts with instructor feedback and had three sources of feedback from which to prepare for their speaking tests (See Figure1).

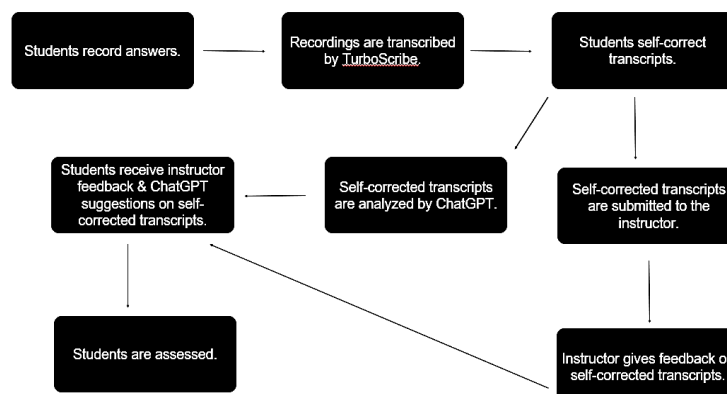


Figure 11. The transcription-task cycle.

For lower-intermediate students, Part III of the IELTS Speaking test was the most challenging. They often had trouble formulating their answers even when they could understand and answer the questions in L1. For these students, the TurboScribe translation function was employed (See Figure 2). They would first answer the question in Japanese, then transcribe their answers from Japanese to English. TurboScribe can translate speech recorded in one language to another language before transcription. After studying the translated transcript, they were asked to make a second recording in English and follow the same steps as the regular transcription-cycle. It is important to give students enough time to analyze the translated transcription to note vocabulary and details. However, they should not be allowed to memorize the answers: they must attempt to answer the question in a similar manner, using their own words and suggestions from the feedback.

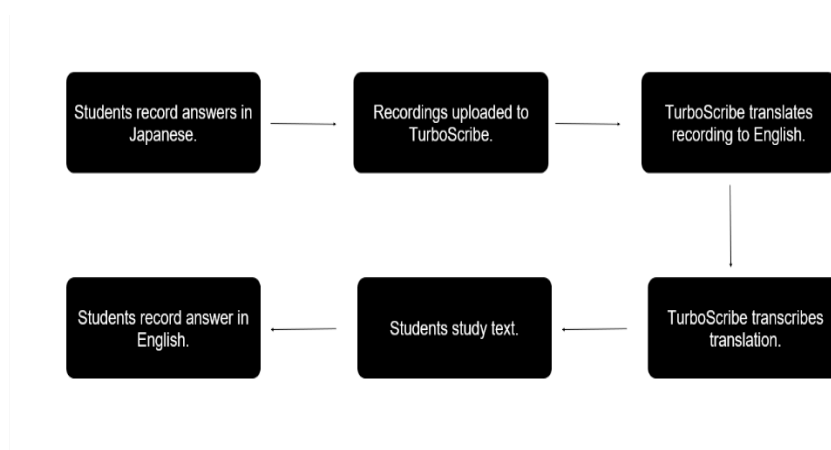


Figure 12. Transcription-task cycle for students answering Part III questions in Japanese.

After a transcription-task cycle, students were assessed using a mix of the same questions from the original set as well as similar questions that they had not seen before.

IV. Assessments

Speaking assessments were carried out after each transcription-task cycle. To avoid memorization of answers, students were assessed immediately after receiving the instructor-feedback. They were allowed a short period of time to study the instructor-feedback and ask questions before the test. Then all students took the speaking test simultaneously by recording their answers to the instructor's questions on their devices. The tests aimed to simulate the IELTS Speaking Test: there were Part I, II, and III questions and the time limits were observed. The recordings were subsequently shared with the instructor through Google Docs. TurboScribe was used to transcribe, the text and was scored and shared with students along with the transcriptions.

Discussion

I. Students' Perceptions

The intervention was carried out in two Speaking & Writing classes: Advanced and Lower Intermediate. A total number of 24 female students participated, 12 in each class. At the end of the academic year, students were surveyed to gauge their perceptions (See Appendix for a list of survey questions). The survey questions and options were provided in both Japanese and English.

Most students agreed that TurboScribe, the error-correction guide, and the instructor's feedback were easy to use and understand (See Figures 3, 4, and 5). However, only 55% of the students said that ChatGPT suggestions were easy to use and understand (See Figure 6). Not surprisingly, the students who found ChatGPT challenging were in the Lower Intermediate group. A closer look at the transcriptions altered by ChatGPT revealed that the suggestions used lower frequency vocabulary, longer sentences with more complex grammatical structures that were probably harder for this group to understand and use for improving their answers. One possible solution that needs to be further explored is how to word the prompts so that lower-level students can also benefit from the ChatGPT advice.

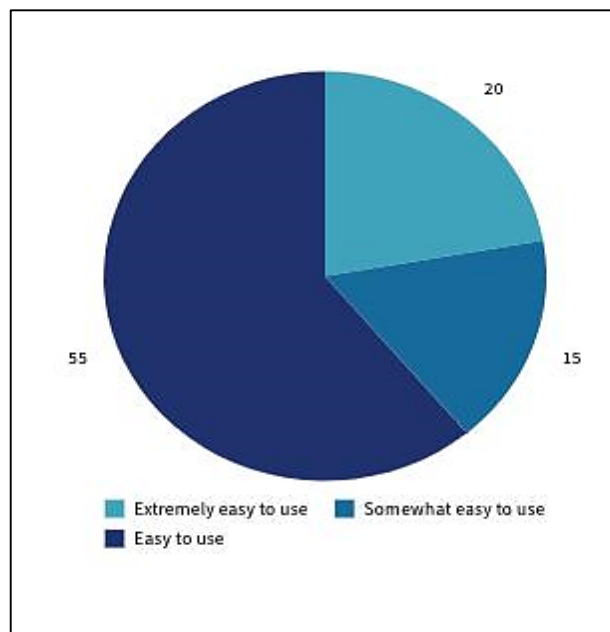
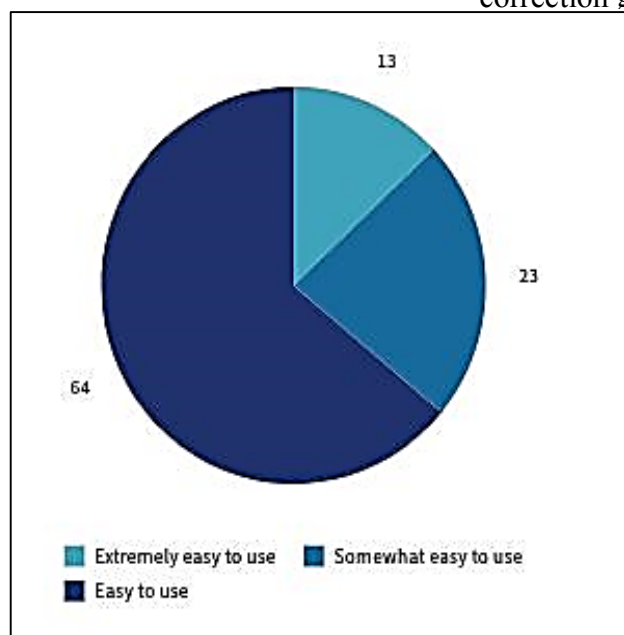


Figure 13. Students' perceptions of TurboScribe (percentages). Figure 14. Students' perception of the error-correction guide (percentages).



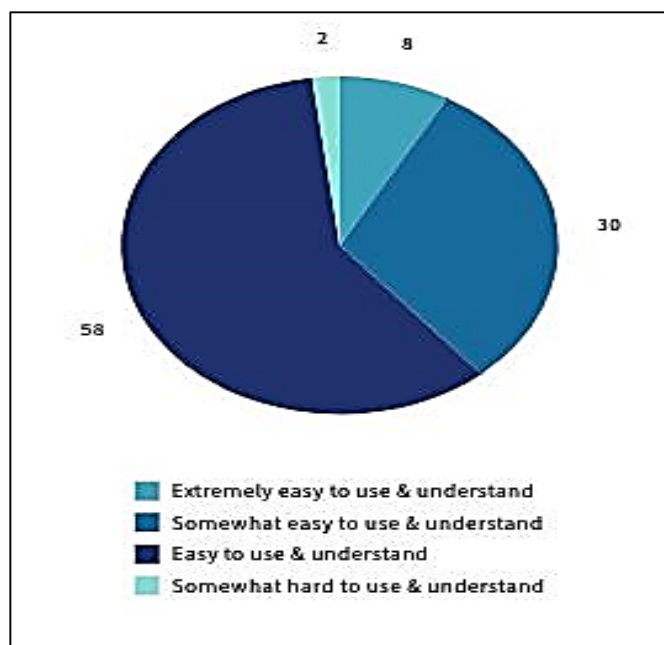


Figure 15. Students' perceptions of instructor feedback(percentages).

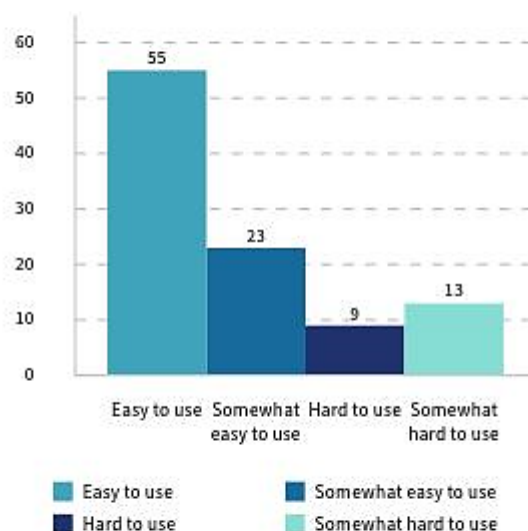


Figure 16. Students' perceptions of ChatGPT (percentages).

Most students seemed satisfied with the AI-supported speaking practice and 64% had made gains of 0.5 to 1 point for their IELTS speaking scores (See Figure 7). These IELTS skills classes are offered once a week and many students have inquired about additional support. Although the university provides an IELTS summer camp for first-year students, many cannot attend due to financial constraints. The department has also arranged for IELTS tutors who can schedule one-on-one free of charge 20-minute sessions, but students often cannot make appointments because of scheduling conflicts with other courses and part-time jobs. It is very likely that the students appreciated the lack of fees and the accessibility of these AI platforms.

Despite this, only 35% said that they were likely to continue to use TurboScribe in the future. This is most likely because after this year none of the students will be taking the IELTS test again, and they no longer need practice. Fifty percent of students said they will continue to use ChatGPT to improve their speaking practice. Most of the positive responses came from the Advanced group. It seems that a lot of these students have been using ChatGPT for grammar study for other courses and they have become accustomed to relying on it.

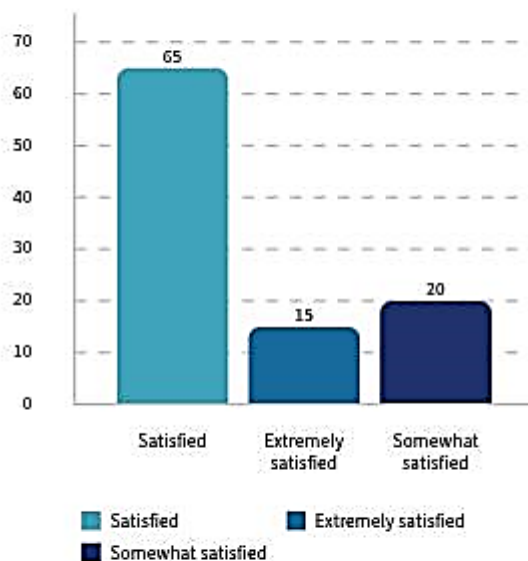


Figure 17. Students' levels of satisfaction with AI-supported speaking practice (percentages).

II. Instructor's Impressions

Listening to audio recordings is very helpful for assessing pronunciation and intonation and it also helps to more easily pinpoint individuals' strengths and weaknesses and tailor advice for improvement. It also saves a lot of class time that can be used for other tasks; instead of using 90 minutes (the standard time for a lesson at this university) to conduct one-on-one speaking tests, an IELTS style assessment with simultaneous group recordings can take just 15 to 20 minutes.

Assessing transcribed speech saves time while ensuring scoring accuracy and fairness. It can also be a good way to collect data and gauge if students are improving over time, what kind of mistakes are repeated, and whether students can successfully incorporate feedback from self-corrections, AI suggestions, and instructor feedback. This data is currently being collected, and results will be disseminated at a future date.

Limitations

Although TurboScribe has been very useful for both learners and instructors, it has some limitations in terms of technological, financial, and academic honesty issues. In terms of technological limitations, there have been issues with the TurboScribe platform claiming that students have used up their daily quota of free translations, even though they

had just registered for the first time ever. This could be due to the fact that since they were all signing up simultaneously from different terminals in the same classroom, the TurboScribe platform might have recognized the IP address and assumed that all the accounts were bots attempting to abuse the system by creating multiple accounts for the same user. To avoid encountering this issue, students should be asked to register from home and do most of the transcription assignments for homework.

A second technological obstacle has been overlapping, inaccurate recordings. For example, if the students are simultaneously recording and sitting in close proximity to each other, it is possible that a fellow student's voice might be picked up by the recorder and transcribed as the device owner's voice. Then the transcription might become jumbled. This has happened when students have made their recordings directly through TurboScribe on a computer or tablet. Although TurboScribe claims that it has a special background noise filtering technology, this still occurs. To avoid encountering this problem, it's a good idea to ask students to sit farther apart from each other and to use their mobile phones to make recordings using a separate recording application on their phone. The recording can then be uploaded to TurboScribe for transcription.

In terms of potential financial issues, it could be costly to register for an unlimited transcription account for instructors who are not subsidized for educational technology expenditures by their institutions. Although the students can use the free TurboScribe accounts for homework assignments, despite the limit of three free 90 minute per day transcriptions, this is not feasible for instructors who need to transcribe a whole class worth of recordings for an assessment.

There are also issues of academic honesty that could potentially arise. Some students have tried to bypass the self-correction stage by asking ChatGPT directly for suggestions and submitting these as their own attempts. Fortunately, it is fairly obvious if intermediate students are cheating in this manner, because the advice from ChatGPT is clearly beyond what they would be able to produce on their own. One way to avoid this problem is to have students print the transcripts in class and monitor while they self-correct. Instructors could also keep copies of shared recordings from students on Google Drive and these can be used to verify suspicious assignments by requiring the student to self-correct during class and compare it with the suspicious assignment.

This case study was conducted with two small groups of students and the data on successful revisions and speaking accuracy and fluency progress needs to be analyzed further before making any recommendations for this practice.

Conclusion

TurboScribe can be a very useful pedagogical and self-study tool despite its limitations. Transcriptions of recordings for speaking tests can make assessments more accurate and save class time that can be dedicated to other tasks. Students seem to view AI-supported transcription in a positive light and can easily adapt to using it both at home and in the classroom to enhance their speaking practice for the IELTS test. Most students' scores for the Speaking component of the IELTS have increased, but without further analysis it cannot be determined if this is due to this intervention. Future studies need to investigate how students actually incorporate the corrections into their speech and which corrections are more useful and accessible: self-corrections, instructor or AI feedback.

References

- Akhter, S., Haidov, R., Rana, A. M., & Qureshi, A. H. (2021). Exploring the significance of speaking skill for EFL learners. *PalArch's Journal of Archaeology of Egypt/Egyptology*, 17(9), 6019–6030. Exploring the Significance of Speaking Skill for EFL Learners
- Dennis, N.K. (2024). Using AI-Powered Speech Recognition Technology to Improve English Pronunciation and Speaking Skills. *IAFOR Journal of Education: Technology in Education*, 12(2), 107-126.
- Dong, L., Jamal Mohammed, S., Ahmed Abdel-Al Ibrahim, K., & Rezai, A. (2022). Fostering EFL learners' motivation, anxiety, and self-efficacy through computer-assisted language learning- and mobile-assisted language learning-based instructions. *Frontiers in Psychology*, 13, 1–15. <https://doi.org/10.3389/fpsyg.2022.899557>
- Harumi, S. 2011. 'Classroom silence: voices from Japanese EFL learners'. *ELT Journal* 65(3), 260–9
- Hirai, A., Kovalyova, A. (2024). Speech-to-Text Applications' Accuracy in English Language Learners' Speech Transcription. *Language Learning & Technology*, 28(1), 3-25.
- Kargar Behbahani, H., Namaziandost, E., & YarAhmadi, M. (2024). The effect of 4/3/2 technique on Iranian EFL learners' speaking fluency: The moderating role of working memory. *Journal of English Language Pedagogy and Practice*, 16(37), 1–24.
- Kim, S. (2024). Reflections on translanguaging as a pedagogical framework in Japanese higher education. *Language Awareness*, 34(1), 1-18.
- Leeming, P., Vitta, J.P., Hiver, P., Hicks, D., McLean, S., Nicklin, C. (2024) Willingness to Communicate, Speaking Self-Efficacy, and Perceived Communicative Competence as Predictors of Second Language Spoken Task Production, *Language Learning*, 74(4), 917-949.
- Lynch, T. (2007). Learning from the transcriptions of an oral communication task. *ELT Journal* 61(4), 311–20.
- Mennim, P. (2003). Rehearsed oral L2 output and reactive focus on form. *ELT Journal* 57(2), 130–8.
- Mennim, P. (2007). Long term effects of noticing on oral output. *Language Teaching Research* 11(3), 265–80.
- Sakui, K. (2004). Wearing two pairs of shoes: language teaching in Japan. *ELT Journal* 58(2), 155–63.
- Stillwell, C., B. Curabba, K. Alexander, A. Kidd, E. Kim, P. Stone, and C. Wyle. (2010). Students transcribing tasks: noticing fluency, accuracy, and complexity. *ELT Journal* 64(4), 445–55.
- Stones, T.P. (2013). Transcription and the IELTS speaking test: facilitating development. *ELT Journal*, 67(1), 20–30. <https://doi.org/10.1093/elt/ccs051>

Appendix

Survey Questions

1. Was TurboScribe easy to use?
Extremely easy to use
Easy to use
Somewhat easy to use
Somewhat hard to use
Hard to use
Very hard to use
2. Were the error-correction guides easy to use and understand?
Extremely easy to use and understand
Easy to use and understand
Somewhat easy to use and understand
Somewhat hard to use and understand
Hard to use and understand
Very hard to use and understand
3. Were the ChatGPT suggestions easy to use and understand?
Extremely easy to use and understand
Easy to use and understand
Somewhat easy to use and understand
Somewhat hard to use and understand
Hard to use and understand
Very hard to use and understand
4. Was the instructor's feedback easy to use and understand?
Extremely easy to use and understand
Easy to use and understand
Somewhat easy to use and understand
Somewhat hard to use and understand
Hard to use and understand
Very hard to use and understand
5. How satisfied do you feel with your AI-supported practice for the Speaking tests?
Extremely satisfied
Satisfied
Somewhat satisfied
Somewhat dissatisfied
Dissatisfied
Extremely dissatisfied
6. Has your IELTS speaking score improved? By how much?
Yes No Current score: _____ Previous score: _____
7. Would you like to continue using TurboScribe to practice speaking English?
Yes No

8. Would you like to continue using ChatGPT to ask for advice on your speaking practice?
Yes No