

# Second Language Learners' Trust in AI Feedback: An Empirical Study Based on the Current Application of AI Technology in Second Language Learning

Xinyi Wang\*

Faculty of Arts, The University of Hong Kong, Hong Kong, China

**Abstract.** This study investigated the level of trust in artificial intelligence feedback among second language learners and its influencing factors, involving 50 undergraduate students majoring in Linguistics at a university in China. It was found that learners generally trust AI feedback, with the degree of trust influenced by language learning experience and the quality of AI tools. Learners with higher levels of trust exhibited greater frequency in using AI tools, more active learning engagement, and lower levels of anxiety. The study suggested enhancing learner trust by improving AI tool quality and developing personalised feedback mechanisms. These findings offer valuable insights for the development of AI educational tools and provide practical recommendations for second language teaching practice.

## 1 Introduction

With the popularization of artificial intelligence in second language learning, various learning platforms and feedback types have been applied in second language learning. These feedback methods are different from traditional second language teacher-student teaching methods and feedback methods in many aspects.

There are various types of feedback in second language learning, such as artificial intelligence grammar correction, vocabulary suggestions and pronunciation evaluation. Based on the appearance of intelligent learning tools, real-time and personalized feedback can be provided to second language learners to help them improve their second language vocabulary, grammar and pronunciation in an effective way [1]. In terms of grammar correction, personalized grammar and spelling checks can be provided according to the needs of users and learners' level of understanding, and more effective corrections can be made. Speech recognition technology can be used to evaluate pronunciation and fluency. Automatic speech recognition applications can help improve students' pronunciation by providing real-time feedback, which can help learners correct their pronunciation and become more accurate in a timely manner [2, 3].

Compared with the feedback from traditional teachers, there are many differences between AI feedback and teacher feedback in terms of speed and personalization. AI systems

---

\* Corresponding author: [U3643755@connect.hku.hk](mailto:U3643755@connect.hku.hk)

can process a large amount of data in a short time and provide real-time feedback based on individual learner data and personalized feedback. On top of that, AI systems can meet diverse learning needs, and therefore, the learning and teaching plans can be optimized to promote effective and personalized learning and teaching. By contrast, traditional feedback methods are often slow and generic. It is difficult to meet various learning situations and personalized learning needs and to develop personalized feedback [2].

With the rapid development of artificial intelligence, more and more second language learners use artificial intelligence technology in their second language learning. People use Grammarly and ChatGPT to help their second language writing learning, and even Duolingo uses it as a feedback tool. This feedback is immediate and personalized. But personalized feedback also has advantages, artificial intelligence may have some problems, such as accuracy and cultural sensitivity. Trust in artificial intelligence feedback becomes the advantages and disadvantages of artificial intelligence feedback. This study attempts to make up for the deficiency of a few empirical studies on the level of trustworthiness that second language learners have in AI feedback and provide theoretical support for designing AI-based learning tools to improve the effectiveness of AI-assisted language learning [4].

## **2 Concept of trustworthiness and its influencing factors**

Artificial intelligence has been successfully applied and used in educational settings. And trust is one of the most important factors for the successful implementation and use of artificial intelligence systems in educational settings. Trust can reduce the uncertainty and risk associated with the adoption of new technology. Trust can be conceptualized in many ways, and in the field of educational technology, it can be explored through concepts such as the Technology Acceptance Model and perceived credibility [5].

Transparency, accuracy, and user experience are influencing factors for students' trust in AI feedback and information from educational technology. Most of the new AI technologies are black boxes, which means that AI systems are not explainable, and AI systems make decisions that learners cannot understand. This phenomenon may make users feel that they are not trusted because they do not know how the AI system makes decisions. Accuracy is one of the influencing factors for learners' trust in AI feedback, AI may generate answers that seem reasonable but are actually wrong or nonsensical, and generate answers with biased content, which will cause learners' trust in AI feedback to decrease [6,7]. In addition, Tiwari also discussed the ethical issues of AI bias in education. In order to gain and maintain trust in AI, it is important that AI systems are fair and accountable in their decision-making processes [7].

## **3 Literature review**

In an experiment to explore students' perceptions of trust in different feedback formats, we randomly chose 91 undergraduate students from UK universities to study the differences in students' perceptions of trust toward three types of feedback: feedback generated by AI, written by human teaching assistants, and optimised feedback written by human-AI collaborative teams. The study used a two-stage design of blind evaluation and informed re-evaluation to collect students' attitude changes in different situations, ensuring that enough students' attitudes changed for analysis. The experimental results showed that in terms of usefulness and objectivity, students considered AI-generated feedback and human-AI collaborative optimised feedback to be more useful and objective than feedback written by human teaching assistants. When students were informed of the source of feedback, the scores of the sincerity of feedback generated by artificial intelligence dropped dramatically.

This shows that students have certain levels of distrust and uncertainty toward feedback generated by artificial intelligence. In contrast, the authenticity scores of human-machine collaboration feedback did not change significantly and were even higher than feedback written by human teaching assistants [8].

In addition, students' experience of using artificial intelligence and gender may also affect their perception of trust in feedback. In the same educational context and situation, students who had more experience using artificial intelligence not only could accurately distinguish feedback generated by artificial intelligence but also showed high levels of trust in the three types of feedback. Students who had more experience using AI actually rated the usefulness and credibility of AI-generated feedback lower. In addition, gender also affects students' perception of trust in different types of feedback. We found that female and non-binary gender students scored more usefully compared with male students. The gender differences in students may be due to the fact that female and non-binary students are more willing to use new technologies and are more likely to consider AI feedback as a useful learning resource. They may also have different expectations and perceptions of the quality of feedback, which affects their overall trust in AI-generated feedback [8].

Although there has been an increasing number of studies on the application of AI in second language learning and teaching, few studies have investigated learners' trust in AI feedback. Firstly, most of the previous studies have concentrated on the usability of AI and its influence on learning outcomes. There is a lack of research on learners' trust in AI. Secondly, there are cultural differences in the research samples. Most of the research samples are Western learners. It is unclear whether there are any significant differences in trust levels of learners from different cultural backgrounds. For instance, Herdiani et al. applied a trust model in an AI-based educational support system [5]. The result showed that it was very important to consider the cultural factor in the process of building trust. However, similar studies on non-Western learners, such as Chinese learners, are still very rare.

To bridge these gaps, the following approaches will be taken by this study. Firstly, the value of trustworthiness will be placed in the research and an attempt to get a further understanding of the role of AI in second language learning and teaching. Secondly, questionnaire survey and open-ended questions will be used to collect multi-dimensional responses of learners' trust in AI feedback. Lastly, the study will be conducted on Chinese learners and attempt to explore the differences in trustworthiness of Chinese learners in different groups.

The findings of this study are expected to provide solid theoretical support for designing more transparent, explainable AI tools and improving the user experience of AI feedback in education. Besides, it will also provide recommendations for second language teaching and help educators to guide second language learning to use AI tools critically and effectively.

## **4 Method**

### **4.1 Research question**

- To what extent do second language learners trust language feedback provided by AI?
- What influences their trust in AI feedback?
- How does trust influence learners' usage behaviour and learning outcomes?

### **4.2 Research design**

A mixed research method is used in this study. Quantitative research is the main body of this study and qualitative research is used as a supplement. Quantitative research is used to collect

data. A questionnaire is used to collect data. The questionnaire is designed to collect data in multiple dimensions. Multiple aspects are designed to collect data. For example, basic information of participants, adopters' usage of AI in second language learning, adopters' trust in AI usage, how trust influences learners' learning behaviour and learning outcomes. It will help to understand second language learners' trust in AI feedback in English writing and its influencing factors. Qualitative research is used to ask open-ended questions. It allows learners to express their views and feelings about AI feedback freely. Through open-ended questions, the reason and motivation of the data can be understood. It offers more background information for the data of quantitative research.

The primary data collection tool of this study is a questionnaire. The questionnaire uses a Likert scale to measure learners' trust in AI feedback. From 1 to 5 represent strongly disagree, disagree, neutral, agree and strongly agree [9]. The questions in the questionnaire are rewritten from the questionnaire used in Wang et al. [10]. For example, the questionnaire has questions like 'I believe AI feedback is reliable' and 'I trust the writing suggestions provided by AI tools'. Participants score their answer on the scale from 1 to 5 based on their personal opinion. In addition to that, multiple questions are designed to collect data on factors that might influence learners' trust in AI feedback. For example, language proficiency, how often they use AI tools and what types of AI tools they use.

### **4.3 Participations**

The sample consisted of 50 university students in China majoring in linguistics. Participants learn English as a second language and use AI tools at least once in the past year. Participants were told before taking part in this study that the questionnaire they filled in would be used for research, but would not be used to intrude on their privacy. Their participation in the study was voluntary.

### **4.4 Data analysis**

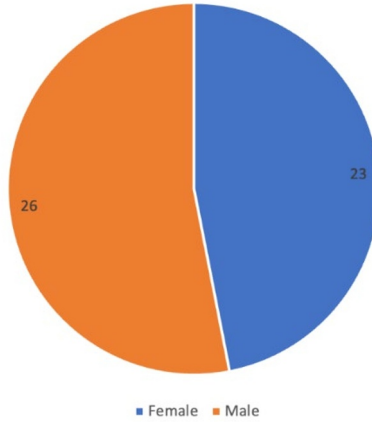
The data analysis of this study mainly used SPSS analytical method. In the process of data analysis, the questionnaire data collected were first organized and pre-processed, such as invalid questionnaire elimination and missing data processing. Then descriptive statistical analysis is used to describe the distribution of the language proficiency level of learners and the number of times of using AI tools. Correlation analysis is then used to explore the relationship between different variables, such as the correlation between the language proficiency level of learners and the degree of trust in the feedback of AI.

## **5 Result and discussion**

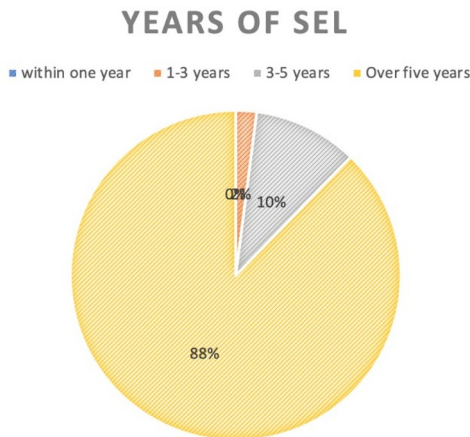
### **5.1 Basic information**

A total of 50 questionnaires were collected, and a total of 49 valid questionnaires were ultimately obtained after collation and pre-processing. As shown in Figure 1, there are 23 males and 26 females. See Figure 2, forty-three of the second language learned for more than five years, one has learned for 1-3 years, and five have learned for 3-5 years. As Figure 3 shows, most of the second language learners who used AI-assisted tools for second language learning used them for 1-2 times a week or 3-5 times a week at the frequency, and there was only one person who used it every day. Twenty-eight used AI tools for writing correction, 20 for vocabulary learning, and 18 for grammar correction. The following is the application of

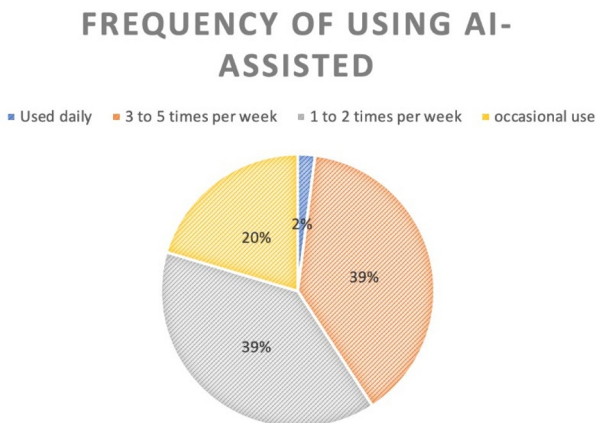
speech recognition and pronunciation correction, as well as spoken dialogue practice. The top 4 tools used by participants were Duoban, Grammarly, Deepseek and Duolingo.



**Figure 1** Participants' gender.



**Figure 2** Years of second language learning.



**Figure 3.** Frequency of using AI-assisted.

## 5.2 AI trust levels and influencing factors

Descriptive statistical analysis shows that second language learners show moderately high levels of trust in overall AI feedback. From Table 1 can be seen that, the mean value of the questionnaire is 4.05, which indicates that most of the learners have some confidence in the feedback of AI. Specifically, the statement 'I trust AI feedback' has a mean score of 4.469 and a standard deviation of 0.71. The results show that most of the learners trust the feedback of AI, but the degree of trust is different.

**Table 1.** Data on trust in AI feedback.

| Question Number | Min | Max | Average | Standard | Median |
|-----------------|-----|-----|---------|----------|--------|
| 7               | 3   | 5   | 4.469   | 0.71     | 5      |
| 8               | 2   | 5   | 4.143   | 0.791    | 4      |
| 9               | 1   | 5   | 2.776   | 1.085    | 3      |
| 10              | 3   | 5   | 4.204   | 0.539    | 4      |
| 11              | 2   | 5   | 3.959   | 0.841    | 4      |
| 13              | 3   | 5   | 4.143   | 0.791    | 4      |
| 14              | 2   | 5   | 3.857   | 0.89     | 4      |
| 14              | 3   | 5   | 4.286   | 0.612    | 4      |
| 15              | 3   | 5   | 4.286   | 0.54     | 4      |
| 16              | 3   | 5   | 4.388   | 0.64     | 4      |
|                 |     |     | 4.0511  |          |        |

Factor analysis found that there were two influencing factors in the factors that affected the degree of trust in the feedback of AI: language learning experience (experience of learning second language and the number of second language learning times, the number of times of using AI) and quality of AI tools (accuracy, speed, interface design and content richness of AI feedback).

The moderately high level of trust means that they can use AI tools as supplementary learning resources. On the other hand, the differences in the level of trust also mean that personalised approaches are needed to improve the level of trust of different learners. The discovery of these two factors through factor analysis provides the basis for the improvement proposals. For example, if the level of trust of learners is improved by enhancing the quality of AI tools (e.g. accuracy, interface design and content richness), then the trust of learners will also increase.

## 5.3 Learning outcomes and behaviour

As shown in Table 2, the correlation coefficients of the first five variables are 0.342, 0.491, 0.400, 0.402 and 0.402 respectively, and all p-values are smaller than 0.05. The correlation coefficient of reducing learning anxiety is 0.108, and the p-value is larger than 0.005. It means that learners' trust in AI feedback has a positive correlation with future increased use

of AI tools, more active participation in the second language learning activity, more willingness to try out new learning features, more willingness to share their learning experiences with others and effective improvement in the second language learning outcomes. However, it has a negative correlation with reducing learning anxiety.

**Table 2.** The correlation coefficients for the first five variables.

| Question Number  | 22     | 23      | 24      | 25      | 26      | 27    |
|------------------|--------|---------|---------|---------|---------|-------|
| 8                | 0.342* | 0.491** | 0.400** | 0.402** | 0.402** | 0.108 |
| *p<0.05 **p<0.01 |        |         |         |         |         |       |

The correlation between trust and usage behaviour and learning outcomes further show that trust in AI is important for learners to use AI tools as auxiliary learning resources in the second language learning process. Learners who trust in the feedback given by AI tools are more likely to use AI tools more frequently, share their learning experiences with others and achieve better learning outcomes. Furthermore, the negative correlation between trust and learning anxiety shows that AI tools can provide a more relaxed learning environment compared with traditional teaching. It is especially beneficial for those second language learners facing anxiety and other challenges in the second language learning. It can improve their learning outcomes in the long term.

## 6 Conclusion

In summary, this study has explored empirically second language learners' trust in AI feedback and found two influencing factors of trust in AI feedback. The findings revealed that second language learners generally maintain a moderately high level of trust in AI feedback and that the level of trust in AI feedback has a positive impact on their usage behaviour and learning outcomes. This study found that the language learning experience of second language learners and the quality of AI tools are two significant influencing factors of trust. Developers of artificial intelligence technologies can enhance the level of trust of learners and optimise the overall learning experience by improving the accuracy, interface design and content richness of AI tools. Furthermore, the findings also show that the level of trust is negatively correlated with learning anxiety. It shows that AI tools can provide a more relaxed learning environment compared with traditional teaching methods. It provides new approaches for second language learners who are facing second language acquisition anxiety. This study is a short-term study, so we could not explore the long-term effects of AI tools on the second language learners. It is necessary for future research to explore the level of trust in different cultures and the long-term effects of AI feedback on learning outcomes. This study provides theoretical basis for the development of AI educational tools. The findings of this study also provide practical suggestions for second language teaching practices, which can help second language teachers effectively promote the use of AI technology.

## References

1. K. Lazzat, The Evolution of Language Learning: Exploring AI'S Impact on Teaching English As a Second Language. *Eurasian Sci. Rev.* **2**, 134-137 (2024).
2. S.S. Evenddy, Investigating AI's Automated Feedback in English Language Learning. *Foreign Lang. Instr. Probe* **3**, 76-87 (2024).

3. A. Mohebbi, Enabling learner independence and self-regulation in language education using AI tools: a systematic review. *Cogent Educ.* 12 (2024).
4. H. Lee and J.H. Lee, The effects of AI-guided individualized language learning: A meta-analysis. *Lang. Learn. Technol.* **28**, 134–162 (2024).
5. A. Herdiani, D. Mahayana, and Y. Rosmansyah, Building Trust in an Artificial Intelligence-Based Educational Support System: A Narrative Review. *J. Sosioteknol.* 23, 101-119 (2024). <https://journals.itb.ac.id/index.php/sostek/index>
6. D. Shin, User Perceptions of Algorithmic Decisions in the Personalized AI System: Perceptual Evaluation of Fairness, Accountability, Transparency, and Explainability. *J. Broadcast. Electron. Media* **64**, 541–565 (2020).
7. R. Tiwari, Explainable AI (XAI) and its Applications in Building Trust and Understanding in AI Decision Making. *Int. J. Sci. Res. Eng. Manag.* 7 (2023).
8. A. Zhang, Y. Gao, W. Suraworachet, T. Nazaretsky, and M. Cukurova, Evaluating Trust in AI, Human, and Co-produced Feedback Among Undergraduate Students. [arXiv:2504.10961](https://arxiv.org/abs/2504.10961) (2025).
9. I.E. Allen and C.A. Seaman, Likert scales and data analyses. *Qual. Prog.* **40**, 64-65 (2007).
10. Y. Wang, H. Wu, D. Ali, and Y. Wang, Development and Validation of a Generative AI-Assisted Attitude Scale for University English Learners in China. *J. Beijing Second Foreign Stud. Univ.* 30-45 (2024).