Shih, JL. et al. (Eds.) (2023). Proceedings of the 31st International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

# AI-powered Collaborative Activities for Chinese Vocabulary Learning

**Xinyu GUO, Yun WEN***
National Institute of Education, Nanyang Technological University, Singapore
* yun.wen@nie.edu.sg

**Abstract:** In recent years, Artificial Intelligence (AI) has significantly increased in digital second language (L2) learning, particularly in supporting vocabulary acquisition. However, research on how AI might facilitate collaborative vocabulary learning is still in its new stage. This study works on investigating the effectiveness of AI-recommended contexts in fostering collaborative language learning among young learners. The research employed a self-developed AI-empowered Chinese vocabulary learning system called ARCHe, which was implemented in primary schools in Singapore. A mixed-methods case study approach was conducted with the 2nd-grade students who spoke English as their first language. Preliminary findings indicate that learning Chinese with ARCHe effectively enhances the academic performance of young learners, with the AI-empowered self-generated contexts feature exhibiting a positive impact on collaborative language learning performance. The study offers insights into the integration of AI in digital language learning, with the potential to enhance L2 learning outcomes for young learners.

**Keywords:** Artificial intelligence, Collaborative learning, Vocabulary learning, Learner-generated Context

## 1. Introduction

In recent years, there has been an unprecedented increase in the use of Artificial Intelligence (AI) in digital language learning, with the pandemic further accelerating its adoption. AI technologies such as machine learning, automatic speech recognition, and natural language processing (NLP) have been utilized to provide automated feedback on student writing (Wu, Wu & Zhang, 2021; Godwin-Jones, 2022) and personalized reading recommendations (Hsu, Hwang, & Chang, 2013; Xiao & Hu, 2019). In the context of Chinese as an L2, AI has shown significant potential in supporting vocabulary learning through character pronunciation, recognition, and writing (Kuo et al., 2022). While the importance of learning contexts in vocabulary acquisition is widely recognized (Atkinson, 2002; Wen, 2021), AI may facilitate interactions among students, recommends learning resources, and automatically generates contexts to foster collaborative learning. However, there is limited research on investigating the role of AI in enriching contextual information, promoting interactions, so as to enhancing collaborative learning for L2 learners.

An AI-enhanced Chinese vocabulary learning system, ARCHe, was designed and developed by our research team. In the ARCHe-supported collaborative learning activity, group students are instructed to construct sentences using provided scenarios. The system can automatically recommend students' artifacts, by which group students will be triggered to create, review, and improve their artifacts continuously. The purpose of this paper is to investigate whether and how AI-recommended contexts can help to promote young learners' collaborative Chinese vocabulary learning. The research questions are as follows:

1. Did the use of ARCHe help to improve young learners' Chinese character learning?
2. How did AI-recommended contexts promote students' collaborative vocabulary learning?

## 2. AI-powered vocabulary learning

Vocabulary learning, as one of the fundamental language skills has received considerable attention in recent studies on technology-enhanced language learning (Burston, 2015). The advent of mobile devices has provided L2 learners with real-time, convenient, and context-rich learning opportunities, particularly in the realm of vocabulary acquisition. Notably, AI tools have emerged as effective means to enhance vocabulary learning. These tools empower students to streamline their language learning journey by delegating specific tasks and receiving personalized learning experiences tailored to their individual needs and progress (Roxana & Fabián, 2023).Considering the potential of AI language generation systems for vocabulary practice and providing diverse assistance (Woo & Choi, 2021), we hypothesize that incorporating this feature into L2 learning tools can promote collaborative interaction.

## 3. ARCHe Design

### *3.1 ARCHe collaborative learning activities*

The ARCHe system comprises two components: a home-based pre-collaborative activity (asynchronous) and a class-based collaborative activity (synchronous) (Wen, 2022). In the home-based activity, students use the ARCHe learning system to complete assigned foundation and creation tasks. The class-based activity involves scenario-based tagging and artifact construction as a group. During the class-based activity, students work in small groups, labeling points of interest in scenario pictures and generating group artifacts. The system's automatic recommendation feature enriches contexts with self-generated artifacts (Figure 1). As students submit group artifacts, the AI-powered system extracts keywords and recommends related learner-generated artifacts for reference, fostering discussions, peer learning, and continuous refinement and improvement of group artifacts.


*Figure 1.* Self-generated contexts recommendation enabled by ARCHe.

## 4. Methodology

This research employed a mixed methods approach with the objective of examining and contextualizing the effectiveness of AI-powered collaborative activities for Chinese vocabulary learning by enriching self-generated-contexts.

### *4.1 Participants*

Table 1. *Details of Participants*

| Teacher | Class | Number of students | Number of groups | Target groups | Ability level |
|---|---|---|---|---|---|
| T1 | C1 | 20 | 7 | T1G1<br>T1G4 | Lower<br>Higher |
| T2 | C2 | 17 | 6 | T2G1<br>T2G5 | Higher<br>Lower |

The study was conducted in two 2nd-grade classes at School A, with a total of 37 student participants. Among them, we selected two groups from each class as the target groups.

### *4.2 Intervention Procedure*

Each class participated in the project for one semester, engaging in class-based collaborative learning activities conducted once every two weeks. Before each class-based session, students were required to complete specific home-based learning tasks. During these class-based collaborative sessions, teachers summarized the performance of the home-based tasks and initially assigned group tasks. With the assistance of the teachers, each group completed a scenario-based tagging and sentence-making task, subsequently reviewed the work of other groups, and provided comments and ratings. The teacher circulated as a facilitator and collaborator throughout the learning process.

### *4.3 Data Source and Analysis*

In this mixed methods study, quantitative data included pre- and post-tests, while qualitative data was collected through various means. Screen recordings, accompanied by audio feeds of student discussions, were captured on mobile devices within each group. Additionally, in-depth structured post-interviews were conducted with teachers and focus groups to provide qualitative insights for in-depth analysis and to interpret the quantitative data. The focus group interviews with young students covered three main areas: group learning task completion, learning mode preferences, and collaboration satisfaction with ARCHe. Moreover, post-interviews were conducted with teachers to gather feedback on teaching using ARCHe and to capture their overall user experiences.

## 5. Findings

### *5.1 The use of ARCHe helped to improve learners' Chinese character learning*

The comparison of the pre & post-tests demonstrates that ARCHe benefits student learning. A total of 37 students from School A. Class 1 and Class 2 completed all the pre-and post-tests. In Class 1, the mean pre-test score was 11.85 (SD=3.45, n=20) out of a total of 30 points. Following the intervention, the mean post-test scores significantly increased to 16.15 (SD=6.09, n=20). The paired t-test revealed a significant improvement in performance among Class 1 students ($t=3.91$, $p<.001$). Similarly, in Class 2, the mean post-test score was 16.118 (SD=5.86, n=17), which was higher than the mean pre-test score of 12.82 (SD=4.23). The t-test also indicated a significant difference between the two scores ($t=3.18$, $p < 0.05$).

Table 2. *Comparison of pre- and post-tests*

| Class | Test | No. of students | Estimated Marginal Means | SD | *t* | *P* |
|---|---|---|---|---|---|---|
| Class 1 | Pre-test | 20 | 11.850 | 3.453 | 3.857 | <.001* |
| | Post-test | | 16.150 | 6.089 | | |
| Class 2 | Pre-test | 17 | 12.824 | 4.231 | 3.182 | .003* |
| | Post-test | | 16.118 | 5.862 | | |

The results of the group artifacts scores provide evidence that ARCHe's learning activities effectively improved the quality of group artifacts. As shown in Table 2, both Class 1 and Class 2 students showed progress in creating artifacts throughout the intervention process. While comparing language skills, we found that students made more significant progress in the content of artifacts.

### *5.2 The recommendation enabled by ARCHe promoted self-context generation*

To explore how AI-recommended contexts promote students' collaborative vocabulary learning, we examined how group interaction took place with the AI-recommended context. The results revealed that the AI-recommendation effectively motivated young students to

review and learn from their peers' artifacts, so as to promote self-context generation in the collaborative learning activity.

Under the guidance of the teachers, intragroup members began exploring peers' artifacts through the self-generated context recommendation, which contained keywords relevant to their own artifacts. During the first lesson in intervention, Teacher 1 and Teacher 2 as collaborators, actively encouraged every group to utilize the self-generated context recommendation. For example, after group 6 of class 1 completed generating the artifacts "我有一个姐姐 (I have a sister), " Teacher 1 suggested: "Let's look at other sentences about '姐姐(sister).'" Then all group members agreed with this suggestion, reviewed and read out one sentence from the self-generated contexts recommendation: "妹妹拿着东西在房间里玩 (my sister is playing with something in her room)" and appraised that was a good sentence.

To facilitate further discussion, we observed that high-level and low-level learning groups employed different strategies in utilizing the AI-recommended contexts. The verbal exchanges from four individual groups are presented in Table 3. For the low-level groups, the recommendation proved more helpful in providing inspiration for generating additional group artifacts. On the other hand, the high-level group, leveraging their judgment skills based on their language proficiency, preferred to comment on and evaluate their peers' artifacts using the recommendation. Their group work through the recommendation primarily focused on improving existing artifacts rather than generating new ones.

Table 3. *Examples of intragroup interactions took place in 4 target groups when using the AI-recommended contexts*

| Focus group | Level | Group Talk Content | Analysis |
|---|---|---|---|
| T1G1 | Low | (Click" 动物 [Animal]" from self-generated context recommendation)<br>➢ Group member 1: 他用的是 '真奇怪'。<br>[They used 'so weird']<br>➢ Group member 2: 这个动物的句子好像有看过。<br>[I saw this sentence before]<br>➢ Group member 3: 我们可以学这些写。<br>[We can learn from these sentences and write similar one] | All three students of this group were engaged in the talk. They prefer to write more similar sentences by browsing peers' artifacts. |
| T1G4 | High | (Click" 雨伞 [Umbrella]" from self-generated context recommendation)<br>➢ Group member 1: 我们没时间了。<br>[We don't have enough time]<br>➢ Group member 2: 让我们再看看有没有更好的句子<br>[let's see whether others have better sentence.]<br>➢ Group member 1: 我觉得我们的比较好。<br>[I think our sentences are better] | This group reviewed and evaluated peers' artifacts by comparing with their own artifacts, which promote more meaningful discussion during collaboration. |
| T2G1 | Low | (Click" 晚上 [Night]" from self-generated context recommendation)<br>➢ Group member 1: 喔，很多很多句子，等一下我们可以 copy 他们的句子。<br>[wow, lots of sentences. Later we can copy their sentences.]<br>➢ Group member 2,3: 好。<br>[Ok] | Since this group have limited language ability, they tried to find inspiration of artifacts generation by reviewing peers' artifacts. |
| T2G5 | High | (Click" 猴子 [Monkey]","朋友[Friend]","很多[very]" from self-generated context recommendation) | All students in this group enjoyed |

| | | | |
|---|---|---|---|
| | | ➢ Group member 1: 有很多关于 “朋友” 的句子。<br>[Lots of sentences about “friend”]<br>➢ Group member 2: 他们的句子没有我们写得好。<br>[Their sentences are not as good as ours]<br>➢ Group member 1: 但这个句子好。<br>[But this sentence is good]<br>(review and read out lots of sentences)<br>➢ Group member 3: 太多句子了，没时间了，我们不要看了。<br>[Two much sentences, the task time is not enough for us to review more.] | reviewing and commenting peers’ artifacts review and evaluation. By comparing peers’ artifacts, they decided if they take further action in their own artifacts’ improvement. |

In the post-interview, regarding the AI-recommended feature, all students expressed their enjoyment of clicking on labels and reviewing their peers' sentences. The two teachers also agreed that this feature stimulated group discussions, idea sharing, and peer learning, enhancing collaborative learning among young learners.

Furthermore, the enthusiasm for utilizing AI-recommended contexts varied between the focus groups from Class 1 and Class 2. Class 1 students displayed curiosity about their peers' artifacts and enjoyed discussing these sentences within their groups. This observation indicates potential differences in teaching approaches between Teacher 1 and Teacher 2. Post-interview findings revealed that Teacher 1 had a consistent focus on developing students' collaborative learning skills from primary 1 up to the current stage. In contrast, Teacher 2 did not extensively discuss collaborative learning during the interview. These findings underscore the importance of considering students' motivation levels and teachers' technological and pedagogical knowledge when integrating AI into language classrooms.

## 6. Conclusion and limitations

The study demonstrated that the collaborative vocabulary learning activities facilitated by ARCHe effectively promoted young learners' academic performance and showed progress in creating group artifacts throughout the intervention process. The positive feedback received from both students and teachers indicates that young learners enjoyed engaging in collaborative vocabulary learning with ARCHe, and teachers observed improvements in students' collaborative learning performance.

The utilization of AI-recommended contexts promoted young students to review and learn from their peers' artifacts. It also enhanced collaborative performance by stimulating self-context generation by idea sharing and discussion. The teachers' guidance in encouraging students to utilize the recommendation function played a vital role in this process. Additionally, the variation in enthusiasm observed between the focus groups from different classes highlighted the importance of effective teaching pedagogy and the role of teachers in fostering successful collaborative learning experiences.

Moreover, the study revealed that different strategies were employed by high-level and low-level learning groups in utilizing the recommendation, with the focus being on artifact generation for the low-level groups and artifact improvement for the high-level groups. This indicates that the effectiveness of using automatic recommendation to modify group artifacts depends on the learners' level of competence. Compared to have students reviewing the self-generated contexts recommendation sentences during the collaborative learning process by themselves, the part of teacher's comments and summary of all groups’ artifacts were more efficient for helping lower-level students’ artifacts improvement.

There are some limitations to this study. Firstly, due to the page constraints, we were unable to provide extensive details about the students’ group learning processes. Secondly, the

effectiveness of students' character learning was only reported based on students' pre-and post-test scores. Additional details will be reported in our forthcoming studies.

## Acknowledgements

This study is funded by Education Research Funding Programme, National Institute of Education (NIE), Nanyang Technological University, Singapore, with project no. DEV 05/21 WY.

## References

Atkinson, D. (2002). Toward a sociocognitive approach to second language acquisition, *Modern Language Journal, 86*, 525–545.

Burston, J. (2015). Twenty years of MALL project implementation: A meta-analysis of learning outcomes. *ReCALL, 27*(1), 4-20.

Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology, 26*(2), 5–24. http://doi.org/10125/73474

Hsu, C. K., Hwang, G. J., & Chang, C. K. (2013). A personalized recommendation-based mobile learning approach to improving the reading performance of EFL students. *Computers and Education*, *63*, 327–336. https://doi.org/10.1016/j.compedu.2012.12.004

Kuo, Y.-C., Yao, C.-B., & Wu, C.-Y. (2022). A Strategy for Enhancing English Learning Achievement, Based on the Eye-Tracking Technology with Self-Regulated Learning. *Sustainability, 14*(23), 16286. MDPI AG.

Roxana, R. F. D. L. V., & Fabián, G. A. (2023). Exploring the Benefits and Challenges of AI-Language Learning Tools. *International Journal of Social Sciences and Humanities Invention 10(01): 7569-7576, 2023*, *2349–2031*. https://doi.org/10.18535/ijsshi/v10i01.02

Wu, L., Wu, Y., & Zhang, X. (2021). L2 Learner Cognitive Psychological Factors About Artificial Intelligence Writing Corrective Feedback. *English Language Teaching, 14*(10), 70. https://doi.org/10.5539/elt.v14n10p70

Wen, Y. (2021). Augmented reality enhanced cognitive engagement: Designing classroom-based collaborative learning activities for young language learners. *Educational Technology Research & Development, 69*, 843–860

Wen, Y. (2022). Design of an AI-powered seamless vocabulary learning for young learners. In Sridhar Iyer et al. (Eds.), Proceedings of the 30th International Conference on Computers in Education Conference (Vol. 2, pp. 659-661). *Asia-Pacific Society for Computers in Education*. https://icce2022.apsce.net/proceedings/volume2

Woo, J. H., & Choi, H. (2021). Systematic Review for AI-based Language Learning Tools. http://arxiv.org/abs/2111.04455.

Xiao, Y., & Hu, J. (2019). Assessment of Optimal Pedagogical Factors for Canadian ESL Learners' Reading Literacy Through Artificial Intelligence Algorithms. *International Journal of English Linguistics, 9*(4), 1. https://doi.org/10.5539/ijel.v9n4p1