Production-Oriented Talk With Chatbot Facilitates Incidental Vocabulary Learning by Chinese English Learners

LIU Huangmei, ZOU Yongwei

University of Shanghai for Science and Technology, Shanghai, China

Researchers never cease their passion about incidental vocabulary learning (IVL) in the field of second language learning (SLA). Chatting with Chatbot is a holistic progress of both production and perception. However, there remains a lack of studies on the role of Chatbot communication on IVL in the following reading. The present study examined Chatbot chatting's influence on IVL in the following reading about science and technology related materials. 40 college students majored in science and technology participated a reading activity after a production-based talking task with Chatbot. The find of the present study showed a positive function of Chatbot in second language learning in practical condition under the present production-based reading tasks, and suggested that using Chatbot in English learning might improve IVL by lowering readers' mental load to increase learners' attention while in the following reading.

Keywords: incidental vocabulary learning, Chatbot, ChatGPT, production-based task

Introduction

Researchers and practitioners in the field of language acquisition and teaching are engaged in the studies of vocabulary learning. Vocabulary learning is usually divided into two categories: intentional vocabulary learning and incidental vocabulary learning (IVL) (Laufer, 2001). The latter is widely accepted as an important part of vocabulary learning. The present study defines incidental vocabulary learning based on the definition given by Longman Dictionary of Language Teaching and Applied Linguistics as learning vocabulary without the intention to learn it, through interaction, communication activities, or reading for content or pleasure. Researchers never cease their passion about IVL, no matter in the field of first language acquisition (FLA) or second language learning (SLL). Many experimental studies have investigated the effects of factors on IVL of FLA, such as text types, tasks, input enhancements, learner strategies, etc., on IVL. There were experimental studied about the relationship between tasks and IVL in SLA (Laufer, 2005; Wu & Xu, 2006), which proved that different types of tasks, such as translation tasks, repeating tasks and sentence making up tasks, had different effects on IVL in both FLA and SLA. Experiments, which studied the relationship between learners' strategy

---

and IVL (Tekmen & Daloglu, 2006), found that some kinds of learner strategies (e.g., meaning inference) were more efficient than the others (e.g., phonologically practice). Experiments examined the relationship between different kinds of input enhancement (e.g., with glossing for the target words, showing the target words in bold letters) and IVL (Webb, 2007) and found that some kinds of enhancements did have better effects than the others.

The development of ChatGPT in various fields has attracted widespread attention from all sectors of society. GPT, or Generative Pre-trained Transformer, is an AI Chatbot program among multiple AI models of OpenAI, known as an "AI language model". Based on the language model of GPT3.5, it adopts Reinforcement Learning from Human Feedback (RLHF) as its training method. As a result, it can accurately understand human language and generate human-like responses and interactions. Researchers have mixed views on the educational changes and opportunities brought by ChatGPT. Some scholars believe that ChatGPT can address issues such as educational inequality, resource uniformity, and lack of personalization (Farrokhnia, Baggen, Biemans, & Noroozi, 2022). However, other scholars argue that applying ChatGPT in education may bring a series of problems, such as academic ethics, limiting thinking, and increasing the burden of homework grading (Rudolph, Tan, & Tan, 2023). Regardless of whether researchers and educators accept it, the era of AI plus education represented by ChatGPT has begun. Foreign language education, both theoretically and practically, needs to explore AI general models like ChatGPT and conduct a comprehensive analysis of their advantages, disadvantages, opportunities, and risks in application. Meanwhile, there remains a lack of studies on the role of Chatbot communication on IVL in the following reading. Chatbot, such as ChatGPT, is becoming more and more popular in second language learning. Researchers believe that it is a holistic progress with both production and perception, which could support foreign language learners from every aspect in reading, writing, listening, and speaking. Some previous studies had found that better background knowledge (scenario familiarity) could facilitate IVL during narrative reading (Pulido, 2007). Therefore, we hold a strong hypothesis that talking with ChatGPT would also prepare learners with more background knowledge for future IVL through reading. The present study examined Chatbot's influence on IVL. ChatGPT is adopted as the AI Chatbot.

Methods

Participants

40 second-year students in a university of Shanghai are invited as participants, who are all majored for science and technology. All students have pasted CET-4 test, which is a national wide English proficiency test for college students. There were no significant differences at their proficiency level ($p = 0.059$, $t = 1.92$, mean = 426). Half of the participants were experimental group, who would chat with ChatGPT for 20 minute with a same given topic before doing the reading task. The others were control group doing the reading task without any pre-chatting activity.

Procedure

Two science innovation related English passages were given to the participants to read with a vocabulary test. One topic is "The development of AI in the last 10 years"; the other is "The development of aerospace technology". Both passages have 334 words. The AI topic was given to the experimental group before the reading task. Then this group was asked to chat with ChatGPT for 20 minutes. They can talk freely with the Chatbot, but the talk had to be restricted to be related to the given topic. After the Chatbot phase, all participants

were asked to read the two passages, and finish the vocabulary test. The vocabulary test included a randomly listed 40 binary choices about 40 words. The test for the vocabulary is not told till all participants finished the reading.

Scoring was done carefully with the help of calculating technique, the equation below.

$$S = (p \text{ hit} - p \text{ false}) / (1 - p \text{ false})$$

Such a calculating technique was adopted to compute the actual performance of word-form recognition. This calculating method was widely used to measure recognition memory and was valid as a scientific method, which can correct the guessing (Pulido, 2007). The hit rate showed the probability of saying "Yes" to the eight target words, while the false alarm rate showed the probability of saying "Yes" to the 16 non-appeared words in the list. All the statistics were analyzed with SPSS 11.5. After the word list test, participants are invited to join in an interview about their feel during the whole procedure. The interview aimed to reflect the mental condition and psychological status of the participants during the experiments. All participants should feel free to express. Therefore, the present study only gives three questions to the participants, after which the participants are free to talk anything they want to talk. The three questions are:

1. How did you feel when you were reading?
2. Did you find the talk with the AI Chatbot helped you when you were reading?
3. Did you enjoy the talk with the AI Chatbot?

Besides these, the participants could express whatever they wanted to say.

Results

Results of Incidental Vocabulary Learning

Table 1 shows the descriptive statistics for the two passages vocabulary test separated by participant group. The mean of hit rate and calculated score of accuracy of the pre-chatted passage (AI) is much higher than that of the other passage within the experimental group and than that of its counter group. The other passage (aerospace) is about the same with the two passages of the control group.

Table 1

Descriptive Statistics of Vocabulary Learning Separated Into Four Groups

Passage	The experiment group	The control group
AI (hit mean)	32	24
AI score	0.75	0.33
Aerospace (hit mean)	24	26
Aerospace score	0.33	0.46

The present study also conducted a one-way ANOVA, with Chatbot talking as the independent variable, and with the hit rate and the individual scoring (see the equation of scoring in the previous section) as the dependent variable. Results of the one-way AVOVA also support the observation from the descriptive statistics, with the performance of experiment group in the prepared passage (AI) significantly better than the other passage within the same group ($p = 0.000 < 0.01$). Such a large difference is out of our expectation to some extent, since we did not expect a very strong short improvement in incidental vocabulary. The two passages have the same number of word count, similar word complexity, similar sentence length, and similar sentence complexity. Therefore, it

is assumed that the passage itself should have no effect on the readers. Thus, the reason for the difference should be an effect of the pre-reading talk with the Chatbot. Meanwhile, we also compared the difference of incidental vocabulary learning of the AI passage between the performance of experiment group and that of the control group; the results still highly agree with the finding in the within group comparison ($p = 0.000 < 0.01$). The difference between the AI passage experiment group and the aerospace passage control group is also significant ($p = 0.03 < 0.05$).

The difference between the aerospace passage IVL of the experiment group and that of the control group is no significant ($p = 0.10 > 0.05$). Neither does that between the aerospace passage IVL of the experiment group and the AI passage IVL of the control group ($p = 0.08 > 0.05$). The difference of the two passages' IVL of the control group is also not significant ($p = 0.19 > 0.05$). Therefore, the results of the present experiment seem to give a very clear suggestion that the pre-reading chatting with the Chatbot does have a positive effect to improve the incidental vocabulary learning of the readers.

Reoccurrence

To avoid the effect of word reoccurrence, we calculated words that had shown during the talk of individual participants with the Chatbot and latter included in the test list at the same time. Each case of the re-appearance of the token was counted as one reoccurrence. All the records of the participants' talk with the Chatbot were scanned carefully. Altogether, for all the 20 participants who had joined in the pre-reading talking with the Chatbot, the average reoccurrence is two, with 0 as the minimum and four as the maximum. And there were no reoccurrences for a word to appear more than twice. All the reoccurred words did all get higher recognition rate (hit accuracy). However, this did not necessarily mean that one repetition gets higher accuracy, since there is no evidence that just one more read incidentally could promise a learning activity. What this phenomenon suggests is that a chatting with the Chatbot did arouse bigger interests and attentions from the readers, which increases the possibility for a new word to be learnt incidentally. On the other hand, we also excluded the reoccurred words, and re-did the one-way ANOVA. This time, the differences of IVL between the Chatbot-condition and the other three condition were narrowed; however, they still remained as statistically significant ($p = 0.033$ with aerospace-experiment group; $p = 0.033$ with AI-control group; and $p = 0.041$ with aerospace-control group). This still suggests that the pre-reading talk with the Chatbot facilitates the followed IVL.

After-Test Interview

This is usually the last phase of a clinical experiment on implicit learning, which unveils the subjects' awareness states. It is used to find whether the subjects learned the target implicit features implicitly or explicitly. The measurement should also reflect the mental condition and psychological status of the participants during the experiments.

An after-test interview has been conducted to record readers psychological condition during the task. Please see the interview design in section of Methods. The results showed that the participants enjoyed the pre-reading talk and found it very helpful for the later reading task (see Figure 1). They emphasized that after chatting with ChatGPT: (a) their load from unfamiliar information was relieved to a very large extent; (b) their psychological load of reading science and the technology related material was also lowered; and (c) some of the vocabulary they read had been understood already during they were talking with ChatGPT.
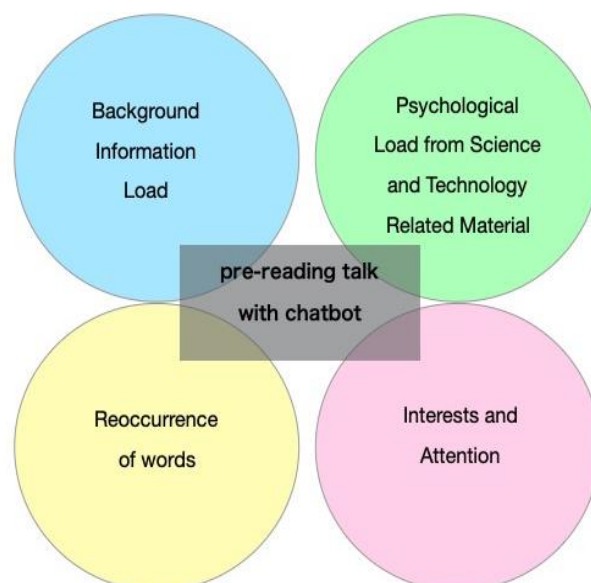
Figure 1. Four highly related effects of pre-reading talk with Chatbot on later IVL.

Conclusion

The present study examined Chatbot's influence on IVL during reading science and technology related materials. In the practical experiment ChatGPT is adopted as the teaching assisting Chatbot, and two passages (one about AI and one about aerospace) are used as the reading materials. Only half of the participants enrolled in a pre-reading talk with ChatGPT on the topic of "AI and human life" for 20 minutes. The major finding of the present study is encouraging, for a condition of a one-time talk with the Chatbot for 20 minutes could improve the incidental vocabulary learning of Chinese English SLL when reading science and technology related materials. Also, the present study finds that reoccurrence of a word from the Chatbot talking could permit readers with much higher rate of IVL. A plausible explanation might be that a chatting with the Chatbot could arouse bigger interests and attentions from the readers, which in return increases the possibility for a new word to be learnt incidentally.

An after-test interview showed that the participants emphasize several helpful points of the pre-reading talk with Chatbot:

1. Their load from unfamiliar information was relieved to a very large extent;

2. Their psychological load of reading science and the technology related material was also lowered;

3. Some of the vocabulary they read had been understood already while they were talking with ChatGPT.

Therefore, pedagogical implications might be gained as that using Chatbot in English learning might be able to lower readers' mental load to increase their attention while reading, which in turn improves IVL. Though limitations, such as participant number and long-term data, existed in the present study due to the restricted experimental condition, it is encouraging to see the positive function of Chatbot in second language vocabulary learning in practical condition under the present production-based reading tasks.

References

Farrokhnia, M., Baggen, Y., Biemans, H., & Noroozi, O. (2022). Bridging the fields of entrepreneurship and education: The role of philosophical perspectives in fostering opportunity identification. *The International Journal of Management Education, 20*(2), 100632. doi.10.1016/j.ijme.2022.100632

Laufer, B. (2001). Reading, word-focused activities and incidental vocabulary acquisition in a second language. *Prospect, 16*(3), 44-54.

Laufer, B. (2005). Focus on form in second language vocabulary acquisition. In S. H. Foster-Cohen, M. P. Garcia-Mayo, and J. Cenoz (Eds.), *EUROSLA yearbooks* (pp. 223-50). Amsterdam: John Benjamins.

Pulido, D. (2007). The effects of topic familiarity and passage sight vocabulary on L2 lexical inferencing and retention through reading. *Applied Linguistics, 28*, 66-68.

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching, 6*(1), 342-363. Retrieved from https://doi.org/10.37074/jalt.2023.6.1.9

Tekmen, E., & Daloglu, A. (2006). An investigation of incidental vocabulary acquisition in relation to learner proficiency level and word frequency. *Foreign Language Annals, 39*, 220-243.

Webb, S. (2007). The effect of repetition on vocabulary knowledge. *Applied Linguistics, 28*, 46-65.

Wu, W., & Xu, H. (2006). The impact of word frequency on incidental learning of English vocabulary. *Journal of Chongqing University (Social Science Edition), 12*(4), 116-121.