https://doi.org/10.29140/97817637116240-11

Using a Custom GPT to Scaffold Corpus-Based Vocabulary Learning for Low-Proficiency Learners

Jing Liu[1,2] and Qing Ma[1*]
[1]Department of Linguistics and Modern Language Studies, The Education University of Hong Kong, Hong Kong SAR, China
[2]School of International Education, Chengdu Polytechnic, Chengdu, China

Abstract

A corpus-based approach is often considered unsuitable for low-proficiency learners, who frequently experience cognitive overload in such learning contexts. To help these learners manage the cognitive demands of corpus-based vocabulary learning, this study designed a custom GPT configured with six scaffolding methods to address three key challenges commonly faced by low-proficiency learners, and integrated it into the framework of corpus-based language pedagogy (CBLP). Unlike teacher-led scaffolding, which may not allow simultaneous support for multiple learners due to inevitable physical constraints, the custom GPT delivers on-demand, contingent support through learner-initiated requests. This mixed-methods study compared two within-group conditions: CBLP with and without custom GPT scaffolding, involving twenty low-proficiency L2 English learners at a vocational college in China. Data sources included vocabulary tests, GPT chat logs, and semi-structured interviews. Results showed that GPT-scaffolded CBLP significantly improved vocabulary knowledge, with the most substantial gains in use. Content analysis of GPT chats revealed that the custom GPT delivered six scaffolding methods in ways that responded contingently to the specific challenges learners faced. Thematic analysis of interview transcripts further indicated that GPT scaffolding demonstrated the characteristic of contingency. Integrating quantitative and qualitative findings, we found that GPT scaffolding appeared

*Corresponding Author E-mail: maqing@eduhk.hk

to have enhanced the effectiveness of CBLP. The present study contributes to the limited but much-needed knowledge of how to integrate contingent scaffolding, enabled by generative AI chatbots, into CBLP lesson designs aimed at enhancing vocabulary learning for low-proficiency learners.

Keywords: Corpus-based language pedagogy (CBLP), GPT scaffolding, contingency, vocabulary knowledge, low-proficiency learners

Background

Despite extensive evidence supporting the effectiveness of corpus-based vocabulary learning (e.g., Karras, 2016; Kaur & Hegelheimer, 2005), a corpus-based approach is often considered unsuitable for low-proficiency learners because heavy cognitive demands may exceed their linguistic competence (Lee et al., 2020). Given this limitation, our study adopted the Corpus-based Language Pedagogy (CBLP) framework (Ma et al., 2022), which emphasizes designing instructional scaffolds to facilitate learners' engagement with corpus-based tasks. However, current corpus instructional practices rely mainly on teacher-led guidance (e.g., Smart, 2014) or pre-designed materials (e.g., Braun, 2005). Although valuable, these supports often fall short in providing contingent scaffolding, meaning moment-by-moment assistance tailored to learners' evolving needs (van de Pol et al., 2010). Drawing on scaffolding theory, we investigated the use of a custom GPT for delivering on-demand assistance during CBLP tasks. Furthermore, prior research on corpus-based vocabulary learning has typically focused on overall vocabulary gains, and few studies have examined its differential effects across individual dimensions. To address this gap, the present study evaluated the impact of GPT-scaffolded CBLP on three dimensions of vocabulary knowledge: form, meaning, and use (Nation, 2022).

Methodology

Designing GPT Scaffolding

We created a custom version of ChatGPT and configured it to implement six scaffolding means as outlined by van de Pol et al. (2010). Aligned with corpus-based activities such as concordance comprehension and inductive discovery, the system delivered three types of support: (1) linguistic support that clarified unfamiliar language items; (2) metalinguistic support that explicated language structures and their functions; and (3) cognitive support that guided learners' lexical inference and pattern detection. Unlike teacher-led scaffolding, which is constrained by the teacher's ability to attend to multiple learners simultaneously in the classroom setting, generative AI tools like the custom GPT offers a promising opportunity to provide on-demand, contingent support in real time (Baidoo-Anu & Ansah, 2023). These three types of support were specifically designed to address the three typical challenges low-proficiency learners face in corpus-based tasks (Liu & Ma, 2025), with the aim of sustaining cognitive engagement with concordance lines and facilitating inductive language discovery.

Participants

The study recruited 20 first-year L2 English learners from a vocational college in China using convenience sampling. All participants scored below 90 out of 150 on the English section of National College Entrance Exam, a criterion used to indicate low proficiency in previous research (Cheng & Liu, 2022) and had a vocabulary size below the 4,000 word-family threshold for intermediate proficiency (Nation & Meara, 2010).

Task Design

In this study, we sampled concordance lines from COCA and SKELL, identifying collocations using a minimum MI score of 3 (Hunston, 2002) to ensure lexical authenticity. Given the potential variability in difficulty across concordance lines (Boulton, 2009), we conducted a lexical coverage analysis to confirm that at least 90% of the text (excluding target words) met the threshold for comprehensibility (Hu & Nation, 2000).

Experiment Design

This study adopted a mixed-methods, repeated-measures design to evaluate the effectiveness of GPT scaffolding in corpus-based vocabulary learning. The intervention included two conditions: one without GPT support and one with GPT support. Participants took part in six 45-minute sessions, with three sessions conducted under each condition. Both conditions followed the same instructional sequence, based on the four-step CBLP lesson design model (Ma et al., 2022), with the only difference being the presence or absence of GPT scaffolding. In the GPT-supported condition, learners used pre-designed prompts to elicit support.

Data Collection and Analysis

A vocabulary pretest and posttest were administered for each intervention condition to assess learning gains. Additionally, semi-structured interviews were conducted at the end of the study to gather qualitative insights into participants' experiences with GPT scaffolding. GPT chat logs were automatically archived during each session for subsequent qualitative analysis. To evaluate vocabulary knowledge gains across the three dimensions both within and between conditions, permutation tests were used to assess statistical significance, given the small sample size and non-normal distribution of the data. Wilcoxon signed-rank tests were used to estimate effect sizes for significant results. Monte Carlo simulations were conducted to assess the statistical power of the findings. Qualitative content analysis was conducted to identify patterns in GPT scaffolding strategies. In addition, thematic analysis of interview data was carried out to uncover key features of GPT scaffolding from the learner perspective.

Results

Within each CBLP condition, a decreasing trend in effect sizes was observed across the three vocabulary dimensions. In the condition without GPT scaffolding, the effect sizes were use ($z = 3.59$, $p < .001$, $r = 0.57$), form ($z = 3.22$, $p = .001$, $r = 0.51$), and meaning ($z = 2.15$, $p = .032$, $r = 0.34$). Similarly, in the GPT-supported condition, the effect sizes were use ($z = 3.54$, $p < .001$, $r = 0.56$), form ($z = 3.48$, $p < .001$, $r = 0.55$), and meaning ($z = 2.67$, $p = .007$, $r = 0.42$). For between-condition comparisons, the use dimension showed significant improvement in favor of the GPT-supported condition ($z = 2.44$, $p = .015$, $r = 0.39$). Effect sizes were interpreted following Plonsky and Oswald's (2014) benchmarks: $r = 0.25$ (small), $r = 0.40$ (medium), and $r = 0.60$ (large). No significant differences were found for form or meaning. Power analyses indicated that, except for meaning, there was sufficient power to detect significant effects for use and form within each condition, as well as between-condition differences in use. In terms of GPT scaffolding strategies, we observed that GPT adapted its responses based on the type of challenge learners faced. Thematic analysis of interview data showed that learners had an overall positive perception of the GPT scaffolding, frequently highlighting its support delivered in a responsive and tailored manner (e.g., "GPT can immediately identify where I'm struggling and provide timely guidance." [S17]).

Discussion

The improvement in the use dimension under the GPT-supported condition may be explained by the nature of corpus-based learning, which offers repeated, contextualized exposure to authentic language, aligning with usage-based theories (Ellis, 2012). GPT scaffolding likely reinforced this effect by directing learners' attention to recurring patterns and providing targeted guidance, thereby activating input enhancement (Sharwood Smith, 1993), promoting noticing (Schmidt, 2001), and increasing involvement load (Laufer and Hulstijn, 2001). As for the form dimension, both conditions showed notable gains, but no significant between-condition differences emerged. This is likely because both groups repeatedly viewed concordance lines with bolded target word forms, which enhanced saliency and triggered implicit learning. Such input flooding (Szczudarski & Carter, 2016) may have already provided sufficient form-focused input, rendering the impact of GPT on this dimension less pronounced. The meaning dimension showed the smallest gains, probably because corpus tasks promote contextual inference that may divert attention from consolidating precise definitions (Tsai, 2019). By integrating quantitative and qualitative findings, the study revealed clear convergence. The custom GPT employed scaffolding strategies that offered contingent support aligned with learners' levels of understanding through learner-initiated interactions. This contingency appears to have contributed to the effectiveness of CBLP, particularly in supporting vocabulary use.

References

Baidoo-Anu, D., & Owusu Ansah, L. (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/10.2139/ssrn.4337484

Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL, 21*(1), 37–54. https://doi.org/10.1017/S0958344009000068

Braun, S. (2005). From pedagogically relevant corpora to authentic language learning contents. *ReCALL*, *17*(1), 47–64. https://doi.org/10.1017/S0958344005000510

Cheng, X., & Liu, Y. (2022). Student engagement with teacher written feedback: Insights from low-proficiency and high-proficiency L2 learners. *System, 109*, 102880. https://doi.org/10.1016/j.system.2022.102880

Ellis, N. C. (2012). Frequency effects. In P. Robinson (Ed.), *The Routledge encyclopedia of second language acquisition* (pp. 260–265). Routledge.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524773

Karras, J. N. (2016). The effects of data-driven learning upon vocabulary acquisition for secondary international school students in Vietnam. *ReCALL, 28*(2), 166–186. https://doi.org/10.1017/S0958344015000154

Kaur, J., & Hegelheimer, V. (2005). ESL students' use of concordance in the transfer of academic word knowledge: An exploratory study. *Computer Assisted Language Learning, 18*(4), 287–310. https://doi.org/10.1080/09588220500280412

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*(1), 1–26. https://doi.org/10.1093/applin/22.1.1

Lee, H., Warschauer, M., & Lee, J. H. (2020). Toward the establishment of a data-driven learning model: Role of learner factors in corpus-based second language vocabulary learning. *The Modern Language Journal, 104*(2), 345–362. https://doi.org/10.1111/modl.12634

Liu, J., & Ma, Q. (2025). Examining corpus-based language pedagogy (CBLP) practices in data-driven learning (DDL) for low-proficiency L2 English learners: A meta-analysis. *Educational Technology & Society, 28*(2), 53–76. https://doi.org/10.30191/ETS.202504_28(2).RP04

Ma, Q., Tang, J., & Lin, S. (2022). The development of corpus-based language pedagogy for TESOL teachers: A two-step training approach facilitated by online collaboration. *Computer Assisted Language Learning*, *35*(9), 2731–2760. https://doi.org/10.1080/09588221.2021.1895225

Nation, P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press.

Nation, P., & Meara, P. (2010). Vocabulary. In N. Schmitt (Ed.), *An introduction to applied linguistics* (pp. 34–50). Hodder Education.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912. https://doi.org/10.1111/lang.12079

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge University Press.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15*(2), 165–179. https://doi.org/10.1017/S0272263100011943

Smart, J. (2014). The role of guided induction in paper-based data-driven learning. *ReCALL, 26*(2), 184–201. https://doi.org/10.1017/S0958344014000081

Szudarski, P., & Carter, R. (2016). The role of input flood and input enhancement in EFL learners' acquisition of collocations. *International Journal of Applied Linguistics, 26*(2), 245–265. https://doi.org/10.1111/ijal.12092

Tsai, K.-J. (2019). Corpora and dictionaries as learning aids: Inductive versus deductive approaches to constructing vocabulary knowledge. *Computer*

Assisted Language Learning, *32*(8), 805–826. https://doi.org/10.1080/09588221.2018.1527366

van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in Teacher–Student Interaction: A Decade of Research. *Educational Psychology Review, 22*(3), 271–296. https://doi.org/10.1007/s10648-010-9127-6