Generative AI tools in CALL: what are the options for teachers and language practitioners?

Branislav Bédi[a], ChatGPT C-LARA-Instance[b], Manny Rayner[b] and Annika Simonsen[c]

[a]Division of Icelandic, Árni Magnússon Institute for Icelandic Studies, ⓘ, branislav.bedi@arnastofnun.is; [b]UniSA STEM, University of South Australia, ⓘ, chatgptclarainstance@proton.me, Manny.Rayner@unisa.edu.au and [c]Department of Icelandic and Comparative Cultural Studies, University of Iceland, ⓘ, ans72@hi.is

Abstract

We present an exploration of generative artificial intelligence's potential in Computer-Assisted Language Learning (CALL), describing recent and expected near-future developments, tools, practical applications, and ethical considerations in the field. We trace different stages of development in CALL and argue that the advent of generative AI has inaugurated a new stage of Intelligent CALL. We consider the possible capabilities and limitations of AI as found in recent studies, compare with our personal, hands-on experience of using generative AI in CALL, and suggest ways to successfully utilise this technology. In particular, we briefly present the C-LARA project, which is a current focus of our activities.

Keywords: artificial intelligence, new periodisation of CALL, teacher training, ethical considerations.

1. Introduction

In this paper, our focus will be to examine the ongoing shift from traditional AI towards the new generative AI paradigm, as it applies to CALL. We offer a brief overview of how recent generative AI technology can be utilised in CALL and outline its capabilities and limitations. Traditionally, AI has been associated with Data-Driven Learning (DDL), corpus research, machine-learning, and areas of language technology including Text-To-Speech (TTS) and Automatic Speech Recognition (ASR): generative AI is rapidly encroaching on all of these, with corresponding consequences for CALL. The rest of the article is structured as follows: a brief description of AI and stages of evolution in CALL; exploring possible capabilities and limitations of generative AI in CALL; and finally, discussion and future directions.

2. A brief evolution of CALL

The evolution of CALL has been marked by distinct stages, with Warschauer (2000) identifying three historical phases: structural (1970s–1980s), communicative (1980s–1990s), and integrative (21st Century). Bax (2003) proposed alternative stages: restricted (1960s-1980s), open (1980s-2000s), and integrated CALL (2000s-present), with the last stage bringing CALL to a so called normalization, where technology becomes ubiquitous (Bax, 2003). This final stage would represent a major transition/paradigm shift (Lv, 2023), rendering technology

2023, Editorial Universitat Politècnica de València

invisible in everyday teaching; this stands in sharp contrast to earlier approaches, where technology is viewed as supporting language education (Nyns, 1989; Shadiev & Wang, 2022), even in advanced applications like personalized learning through data analysis and instant feedback (Kohnke et al., 2023). Traditional AI relies on pre-existing data for pattern recognition, predictions, and classification. In contrast, generative AI also uses its training data to create original content, opening new perspectives for teaching and learning. Developing new digital skills, such as learning how to generate effective prompts (Kohnke et al., 2023), has now become part of Intelligent CALL. Technology has long supported language education and generative AI has brought about a new era, reshaping how we teach and learn in CALL.

3. Generative AI in CALL

This section explores how generative AI systems, e.g. ChatGPT-4[1], Bard[2], DALL-E-2[3], Midjourney[4], and Stable Diffusion[5], can be used for creating multimodal language learning resources. Please note that performance is strongly language dependent. Bard can currently only produce output in few languages. Even ChatGPT is currently much stronger in well-resourced, commercially important languages that have been prioritised by OpenAI. Performance in less-resourced languages is, however, improving quickly as more training data is deployed in training the models.

3.1. Capabilities of generative AI in CALL

With the above caveats, generative AI tools enable fast and relatively reliable ways of generating texts and images; as we write, audio capabilities are just about to be released for ChatGPT[6] and are promised for Google's Gemini[7]. Our observation is that the generative AI's current capability enables CALL teachers and practitioners to create useful tasks that help with creating new knowledge and content. Inspired by the most recent discussion and research (Bell, 2023; Crosthwaite & Baisa, 2023; Dilmegani, 2023; Ray, 2023), potential ways of using generative AI in CALL include the following:

1. Personalising lessons by learners' past performance, language skills, feedback, and learning abilities;
2. Organising, planning, and assessing lessons and course design to enhance individual learning experiences;
3. Creating and generating various learning materials, such as quizzes and flashcards, discussion questions, exercises, reading lists, study guides, summaries, video lecture scripts, language translations, grammatical correction or sentence analysis, and tagging text with parts of speech;
4. Generating new text and multimodal content;
5. Reviving outdated or poor-quality learning materials;
6. Offering tutoring through interactive chatbot discussions (conversational AI);
7. Developing digital skills for effective prompts and collaborative use of generative AI tools;
8. AIs collaborating with both technical and non-technical human experts to enhance CALL platforms; and
9. Integrating generative AI into linguistic and educational tools, supporting data-driven learning in real contexts, and Learning Management Systems (LMSs), transforming CALL to a more intelligent level, e.g. Duolingo Max[8], TinyStories with datasets of short stories created by ChatGPT-3.5 and ChatGPT-

---

[1] https://openai.com

[2] https://bard.google.com/

[3] https://openai.com/dall-e-2

[4] https://www.midjourney.com

[5] https://stablediffusionweb.com

[6] https://openai.com/blog/chatgpt-can-now-see-hear-and-speak

[7] https://lifearchitect.ai/gemini/

[8] https://blog.duolingo.com/duolingo-max/

4[9], C-LARA with integrated ChatGPT-4 functionality to generate texts, translations, and tagging (cf. section 3.3).

Harnessing these and related capabilities, the generative AI technology will soon offer a new generation of more advanced educational tools for language learning which will rapidly become part of CALL.

3.2. Limitations of generative AI in CALL

Although current research describes generative Large Language Models (LLMs) as revolutionary AI-driven technology, which has indeed a great potential in education (Kasneci et al., 2023; Kohnke et al., 2023), the same sources also voice various concerns:

1. As already noted, the fact that there is insufficient training data for all but the most highly prioritised languages means that the quality of generative AI tools is strongly language dependent.
2. Ethical aspects and copyright issues when using AI-generated output are important. The AI generates outputs guided by user prompts, but leveraging styles and genres taken from its training data (books, articles, paintings, audio recordings, videos), particularly if the material has been used without the original authors' consent, raises the possibility that some output can arguably be considered as plagiarised. Human editing of the output may alleviate the problem, but this both detracts from the usefulness of the generative AI technology and will not necessarily be accepted.
3. Other ethical and legal limitations include privacy and security issues (Bozkurt & Sharma, 2023; Kasneci et al., 2023; Kohnke et al., 2023; Short & Short, 2023).
4. Paid access to advanced versions of generative AI tools is another concern because monetisation limits access (Bozkurt et al., 2023). Language institutions and teachers have often limited resources that restrict them from using paid services for teaching and learning.
5. Shortage of training data may be manifested in cultural as well as linguistic problems. We ourselves found in a recent experiment with Icelandic (cf. Section 3.3 below) that cultural aspects from North America were often imported into Icelandic settings.
6. The widely remarked-on phenomenon of confabulations or 'artificial hallucinations' (Alkaissi H., McFarlane, 2023; Short & Short, 2023), endemic to generative AI, means that the factual aspects of generated content can in general not be considered reliable.
7. Lack of the necessary digital skills to generate effective prompts may affect the quality of the output.
8. Inequity and inequality are another issue to consider. Access to and using generative AI tools may be limited due to possible learners' disabilities and impairments (Kasneci et al., 2023).

3.3. Utilizing generative AI tools: our experience

Another study (Simonsen & Bédi, 2023) describes initial experiments with using generative AI tools to create multimodal texts directly. ChatGPT-4 was used to generate 15 anecdotal short stories for L2 Icelandic learners at various levels and ages. We added appropriate images using DALL-E 2, Midjourney, or Stable Diffusion. It was interesting to find that the same prompts used in the three tools gave substantially different results, and that the different AI systems required different amounts of detail in the prompts. In each case, the content creator could fine-tune the prompts by adding or changing wording to reach an image and style they personally liked.

A more substantial effort is currently under way in the shape of the ChatGPT-Based Learning And Reading Assistant (C-LARA; https://www.c-lara.org/; Bédi et al., 2023a, b). C-LARA takes inspiration from the earlier LARA platform (Akhlaghi et al. 2019) but represents a complete rewrite. Similarly to LARA, the basic goal is to develop a flexible platform that allows users to construct multimodal texts useful for L2 learners, but, as the name suggests, ChatGPT-4 stands in the center. The AI is used in two separate and complementary ways, both as

---

[9] Eldan and Li (2023). TinyStories: How Small Can Language Models Be and Still Speak Coherent English?. https://arxiv.org/pdf/2305.07759.pdf .

a software *engineer* and as a software *component*. In its software engineer role, ChatGPT has played a major part in developing the platform, collaborating closely with a human partner to write about 90% of the code and contribute the greater part of the software design.

In its software component role, ChatGPT can perform all the core language processing tasks. It can write short texts in response to user prompts, segment them into sentences and lexical units, annotate them with glosses and lemma/part-of-speech information, and add summaries. The prompts used to perform these operations are constructed from templates combined with few-shot sets of examples, facilitating language-specific tuning. The results are combined with TTS audio (the option of human-recorded audio will soon be added) and can be posted openly in a simple social network where other users are able to leave ratings and comments. The human annotator is able to edit the results of each stage, and there is functionality to compare different versions of the plain and annotated text, making it easy to estimate error rates for the different AI-based operations. For well-resourced languages prioritized by OpenAI, like English and Mandarin, ChatGPT-4 can write fluent, engaging texts in a wide variety of genres with an error rate of well under 1%; accuracy for glossing and tagging for these languages is typically in the mid-single digits, with the majority of the errors related to treatment of multi-word phrases. For less well-resourced languages, error rates can be much higher. A paper describing the issues in more detail, using a sample of five languages, is under review. We expect a public deployment of C-LARA to be available by the time this paper is published.

4. Summary and conclusions

With the utilisation of different AI technologies in CALL, a new era of Intelligent CALL has emerged in language education. We have tried to give a brief overview of the capabilities and limitations of generative AI. Generative AI tools offer many benefits for CALL, most importantly providing swift and reliable means for creation of new content. Potential applications range from personalised lesson planning and content creation to restoration and revival of outdated learning materials and tutoring. However, successful use of these tools typically requires effective input fine-tuning and prompt generation (Bozkurt, 2023; Kohnke et al., 2023; Short & Short, 2023; Kasneci et al, 2023). Use of generative AI in CALL has also raised concerns related to plagiarism, biases, privacy, security, equality, and equity, as well as careful consideration of ethical and copyright issues, and training data limitations mean that the tools so far only work well in the relatively small number of languages prioritised by the manufacturers. In conclusion, we briefly described our experiences with generative AI, in particular the collaborative open source C-LARA platform, where ChatGPT is used both as a software engineer and as a software component.

References

Alkaissi H. & McFarlane, S. I. (2023). Artificial Hallucinations in ChatGPT: Implications in Scientific Writing. *Cureus 15*(2). PMC – PubMed. https://assets.cureus.com/uploads/editorial/pdf/138667/20230219-28928-6kcyip.pdf.

Akhlaghi, E., Bédi, B., Butterweck, M., Chua, C., Gerlach, J., Habibi, H., Ikeda, J., Rayner, M., Sestigiani, S., & Zuckermann, G. (2019). Overview of LARA: a learning and reading assistant. In *Proceedings of SLaTE 2019: 8th ISCA Workshop on Speech and Language Technology in Education*, 99–103. https://doi.org/10.21437/slate.2019-19.

Bax, S. (2003). CALL—past, present and future. *System 3*(1), 13–28. Elsevier. https://doi.org/10.1016/S0346-251X(02)00071-4.

Bell, E. (2023). Generative AI: How It Works, History, and Pros and Cons. Inverstopedia. https://www.investopedia.com/generative-ai-7497939.

Bédi, B., ChatGPT-4, Chiera, B., Chua, C., Ni Chiaráin, N., Rayner, M., Simonsen, A & Zviel-Girshin, R. (2023a). ChatGPT + LARA = C-LARA. Research Gate. https://www.researchgate.net/publication/373952306_ChatGPT_LARA_C-LARA.

Bédi, B., ChatGPT-4, Chiera, B., Chua, C., Ni Chiaráin, N., Cucchiarini, C., Rayner, M., Simonsen, A & Zviel-Girshin, R. (2023b). ChatGPT-Based Learning and Reading Platform: Initial Report. Technical report. ResearchGate.

Bozkurt, A. (2023). Generative artificial intelligence (AI) powered conversational educational agents: The inevitable paradigm shift. *Asian Journal of Distance Education 18*(1), 198-204. Asian Society for Open and Distance Education (ASODE), Japan. http://www.asianjde.com/ojs/index.php/AsianJDE/article/view/718/399.

Bozkurt, A., & Sharma, R. C. (2023). Challenging the status quo and exploring the new boundaries in the age of algorithms: Reimagining the role of generative AI in distance education and online learning. *Asian Journal of Distance Education 18*(21), i-viii. https://doi.org/10.5281/zenodo.7755273.

Bozkurt, A., Xiao, J., Lambert, S., Pazurek, A., Crompton, H., Koseoglu, S., Farrow, R., Bond, M., Nerantzi, C., Honeychurch, S., Bali, M., Dron, J., Mir, K., Stewart, B., Costello, E., Mason, J., Stracke, C. M., Romero-Hall, E., Koutropoulos, A., Toquero, C. M., Singh, L Tlili, A., Lee, K., Nichols, M., Ossiannilsson, E., Brown, M., Irvine, V., Raffaghelli, J. E., Santos-Hermosa, G Farrell, O., Adam, T., Thong, Y. L., Sani-Bozkurt, S., Sharma, R. C., Hrastinski, S., & Jandrić, P. (2023). Speculative futures on ChatGPT and generative artificial intelligence (AI): A collective reflection from the educational landscape. *Asian Journal of Distance Education, 18*(1), 53-130. https://doi.org/10.5281/zenodo.7636568.

Crosthwaite, P. & Baisa, V. (2023). Generative AI and the end of corpus-assisted data-driven learning? Not so fast! *Applied Corpus Linguistics 3*, 1-4. Elsevier. https://doi.org/10.1016/j.acorp.2023.100066.

Dilmegani. C. (2023). op 6 Use Cases of Generative AI in Education. AIMultiple. Last update 22 April 2023. Accessed on 30 July 2023: https://research.aimultiple.com/generative-ai-in-education/.

Eldan, R., & Li, Y. (2023). TinyStories: How Small Can Language Models Be and Still Speak Coherent English?. https://arxiv.org/pdf/2305.07759.pdf.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G.. Michaeli, T., Nerdel, C.. Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J. Kuhn, J., Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences 103*. Elsevier. https://doi.org/10.1016/j.lindif.2023.102274

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for Language Teaching and Learning. *RELC Journal*, *54*(2), 537-550. https://doi.org/10.1177/00336882231162868.

Lv, Z. (2023). Generative artificial intelligence in the metaverse era. *Cognitive Robotics 3*, 208-217. Elsevier https://doi.org/10.1016/j.cogr.2023.06.001.

Nyns, R. R. (1989). Is intelligent computer-assisted language learning possible?. *System 17*(1), 35-47. Elsevier. https://doi.org/10.1016/0346-251X(89)90058-4.

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems 3*, 121-154. Elsevier. https://doi.org/10.1016/j.iotcps.2023.04.003.

Shadiev, R. & Wang, X. (2022). A Review of Research on Technology-Supported Language Learning and 21st Century Skills. *Frontiers in Psychology 13*. Frontiers. https://doi.org/10.3389/fpsyg.2022.897689.

Short, C. E. & Short, J. C. (2023). The artificially intelligent entrepreneur: ChatGPT, prompt engineering, and entrepreneurial rhetoric creation. *Journal of Business Venturing Insights 19*. Elsevier. https://doi.org/10.1016/j.jbvi.2023.e00388.

Simonsen, A. & Bédi, B. (2023). Using Generative AI tools and LARA to create multimodal language learning resources for L2 Icelandic. In *B. Bédi, Y. Choubsaz, K. Friðriksdóttir, A. Gimeno-Sanz, S. Björg Vilhjálmsdóttir & S. Zahova (Eds.), CALL for all Languages - EUROCALL 2023 Short Papers*. https://doi.org/10.4995/EUROCALL2023.2023.16994