




EDITORIAL

Chatgpt in and for second language acquisition: a call for systematic research

ZhaoHong Han 

Teachers College, Columbia University
Email: han@tc.columbia.edu.

(Received 10 January 2024; Revised 10 January 2024; Accepted 19 January 2024)

Applications of Artificial Intelligence (A.I.)-powered tools are by no means new to researchers and practitioners in second or foreign language learning. For one, chatbots, also known as conversation agents (CAs) or conversational A.I., have generated considerable interest in the field of applied linguistics (see, e.g., Fryer et al., 2020; Huang et al., 2022; Ji et al., 2023; Xiao et al., 2023; Zou et al., 2023). CA platforms are language models in that they simulate real-world language use, to a certain extent. A recent scoping review by Xiao et al. (2023) suggests that CA-aided interaction carries the flavor of authenticity, is motivating, enjoyable for language learners, and is used by teachers as a supplementary source of practice for language learners and a tool for formative assessment. This research, according to Xiao et al. (2023), has been dominated by an interest in users'—learners' more than teachers'—perceptions of their experiences, a sign likely of a commercially driven interest as well as an embryonic stage of academic research.

A.I.-powered language models have evolved at a breathtaking pace. The recent release of ChatGPT (Generative Pre-trained Transformer) by OpenAI took the world by storm with its jaw-dropping capabilities, from natural language understanding to conversation generation, language translation, text summarization, grammar correction, paraphrasing, and more. Billed as a large language model, ChatGPT has been trained on an outsize (on the scale of hundreds of gigabytes) and diverse dataset of texts published on the internet on an array of topics, learning the statistical patterns and relationships between words and phrases in text. The end game is a mathematical architecture of hundreds of millions of parameters, resembling a neural network in the human brain capable of performing generative functionalities like the types aforementioned. It is anticipated that future iterations of ChatGPT will see an exponential growth of parameters and, hence, an ever more sophisticated algorithm, which, aided by Herculean computing power, will yield even greater human-like language capabilities (Metz, 2024; Rudolph et al., 2023).

For such a gigantic, actual, and potential computer simulation of natural language usage, societal and academic reactions are mixed. Enthusiasts are exhilarated about ChatGPT's language modeling capabilities. Conversely, critics are gravely concerned that “the eerily humanlike chatbot” (Satariano & Kang, 2023) can do catastrophic harm to humanity. Other than its liability for spreading bias and controversial content, the fear of ChatGPT thwarting human learning potential and taking away human livelihoods looms large, so much so that governments are in disarray as to how to respond in terms

of regulating ChatGPT's development and use (Satariano & Kang, 2023). Notwithstanding, there is a general reckoning that ChatGPT (generative A.I., for that matter) is here to stay, like many technologies, such as word processing, search engines, and social media. In the grand scheme of education, bringing ChatGPT into the classroom has increasingly been viewed as inevitably supporting future workforce development training (The Southeastern Conference, 2023).

As a large language model, ChatGPT drives home a unique array of critical issues for language researchers and educators. Not surprisingly, the reactions here are mixed as well. Some people are outright skeptical, some are sitting on the fence, but many choose to embrace the technological innovation and contemplate how to leverage ChatGPT's capabilities in educating the next generation of foreign language learners (see, e.g., Baskara & Mukarto, 2023; Bonner et al., 2023; Fryer et al., 2020; Hong, 2023; Ji et al., 2023; Kohnke et al., 2023; Li et al., 2023; Waschauer et al., 2023).

Scholarly work, meanwhile, is surfacing in academic journals, much of which involves legislating what to do. Tseng and Waschauer (2023), for example, promulgated a five-part pedagogical framework for educators wishing to engage with ChatGPT: *understand*, *access*, *prompt*, *corroborate*, and *integrate*. This framework is aimed at educating learners about ChatGPT and navigating it as an educational resource while taking into account its limitations. It begins by helping students understand what ChatGPT is and ends with teaching students how to "appropriately and ethically incorporate AI-generated content into their own work, including how to note and cite the role of AI-based tools into their writing process" (p. 261).

Empirical studies on ChatGPT in the context of Second Language Acquisition (SLA), however, remain to be seen. The pre-ChatGPT body of literature, as earlier mentioned, features mostly surveys of learners' perceptions of their experiences with A.I. tools, which barely scratches the surface of substantive issues. As Xiao et al. (2023) pointed out:

Despite the learning benefits of perception research, the existing literature has not systematically examined the actual linguistic improvement of commercial CAs. The most common tasks assigned to learners were to try out the commands, play games and ask the CA questions, which largely reflects the current available features of CAs, only a handful of studies have redesigned the dialogue system to fit into specific learning contexts or redesigned the tasks for specific learning purposes. (p. 12)

It is clear that the potential for CAs has yet to be substantiated, through a systematic agenda of scholarly research.

ChatGPT marks a clear inflection point in the development of general purpose A.I. and high stakes chatbots, presenting an unprecedented need and opportunity to explore and examine its relevance to second language learning. The concerns relevant to SLA are not quite like those facing higher education regarding teaching and assessment, such as the validity of take-home essays, plagiarism, ghost-writing, etc., although foreign and second language educators should be wary of any nefarious consequences of using ChatGPT that may hurt rather than help students to develop thinking skills to function effectively in society.

In SLA, there is a pressing need for bidirectional research into ChatGPT's language learning affordances, defined as potentialities for language development. On the one hand, we need to empirically investigate its alleged functionalities with an eye toward identifying their potential for language learning. On the other hand, we need to find out if learners interacting with ChatGPT actually use the affordances, and if so, how and to

what end. In addition, we must examine if learners, while interacting with ChatGPT, create affordances for themselves. In other words, do any affordances *emerge* as the learner engages with ChatGPT?

The three strands of research each have their own unique questions to address. The first strand concerning the inherent make up and capabilities of ChatGPT should, among many questions, take up the following targeting ChatGPT as a language model:

- What kind of language model is given by ChatGPT?
- What types of language does ChatGPT model?
- What are the model's characteristics, including lexical, morphosyntactic, pragmatic, and discourse attributes?
- To what extent does the model reflect real-world human language use?
- How representative is it of natural language use in the real world?
- What constraints does the model place on language use?
- What is missing from the language model?
- What types of biases might it contain?
- What is artificial about it?
- How rigid is it?
- How capable is ChatGPT of understanding the learner's prompts?
- How adept is it at responding to the learner's needs?
- What kind of learner language prompts may evade ChatGPT?
- How scalable are ChatGPT's capabilities?

Language learning (and language instruction, for that matter) is teleological (Han, Kang, Sok, 2023; van Geert, 2023). Therefore, investigating these questions is fundamental to understanding ChatGPT vis-à-vis language learning and the expected learning outcome. From the perspective of a language learner, ChatGPT models how language is used or should be used.

The second strand of research focuses on what learners do with ChatGPT's potential affordances. Here, research needs to focus on the interaction between ChatGPT and the language learner, exploring its dynamics. Questions of relevance can include but are not limited to the following:

- How does the learner use ChatGPT, and for what purpose?
- What particular functionalities does the learner use?
- Does the interaction lead to any change in the learner's linguistic behavior? If so, what kind of change?
- What level of proficiency is required for the learner to engage with ChatGPT?
- What is the learner's level of engagement with ChatGPT—superficial, substantial, or sophisticated?
- What input does the learner provide to ChatGPT in the form of a prompt?
- How does the learner revise their prompt when ChatGPT cannot comprehend it?
- Does the learner's use of prompts change over time?
- What is the language model the learner builds from their interaction with ChatGPT? Or how does the learner train on ChatGPT's output?
- What is the optimal duration and intensity of interaction to yield changes in the learner's use of their L2?
- How does the learner's interaction with ChatGPT compare to human-human interaction, and how do the two modes of interaction compare in terms of language learning efficacy?

- What neurological networks are activated as learners interact with ChatGPT?
- Do the neuro-networks change over time as the learner's interaction with ChatGPT continues? If so, how?

At its core, the second strand of research seeks to understand (a) the socio-cognitive processes induced by the learner's interaction with ChatGPT and (b) the learner's psycholinguistic processes, including processing of input, that is, ChatGPT's output. How does the learner negotiate their understanding of the input, etc.?

And there is a third strand of research on the SLA agenda. The learner's interaction with ChatGPT, an A.I. chatbot, does involve and require human agency. Thus, the third strand of research is about the learner, and questions of relevance may include the following:

- Does the learner's interaction with ChatGPT change over time? If so, in what ways?
- In the learner's use of a ChatGPT functionality, do they become more skilled?
- Does the learner get more creative than using only the ostensible affordances provided by ChatGPT?
- What does the learner do before and after interaction with ChatGPT?
- Do learners differ in their predisposition to ChatGPT, ability to engage with it, and learn from its output?
- What does the learner learn from ChatGPT?
- Do the learner's emotions fluctuate as they interact with ChatGPT?
- Does their level of engagement fluctuate as they interact with ChatGPT?
- Does the learner's extended interaction with ChatGPT affect their perception of identity, self-esteem, and self-efficacy?
- How does the learning that takes place when interacting with ChatGPT compare to the learning from a human interlocutor (e.g., a teacher)?
- Does the learner's extended interaction with ChatGPT affect their ability, and/or willingness, to interact with a human interlocutor?
- To what extent does the learner's extended interaction influence their writing?
- Does the learner use ChatGPT as a source of feedback? If so, how? And what is their interpretation of the feedback? And how does it influence their subsequent writing, and, by extension, their listening and speaking?
- Does the interaction encourage the learner's creative use of language? Does it stifle it?

In pursuing the three interrelated strands of concern, in particular, strands 2 and 3, researchers would benefit from employing as their conceptual frameworks SLA theories (VanPatten et al., 2020), especially the so-called transition theories that seek to account for second language (L2) development (Gregg, 1996). Examples are Basic Language Cognition Theory (Hulstijn, 2015), sociocultural theory (Lantolf et al., 2020), cognitive-interactionist theory (Gass & Mackey, 2020), usage-based approaches (Ellis & Wulff, 2020), and Complex Dynamic Systems Theory (Larsen-Freeman, 2020). The role of theories can be multi-faceted. It helps frame the study design, formulate research questions, interpret findings, and even unify studies that use similar theoretical frameworks. The latter, among other things, would enable the accumulation of evidence, which is critical to advancing our understanding, from discovery to validation. In addition, for strand 1, it may be instructive to consider the machine learning mechanisms underlying ChatGPT, such as supervised and reinforcement learning (Hong, 2023; Rudolph et al., 2023), and other A.I. applications to the learning sciences.

Just as theories can guide empirical research, studies on ChatGPT-mediated learning can contribute to the refinement of theories. Thus, theory and research form a symbiotic relationship that advances SLA as a scientific discipline.

ChatGPT-mediated language learning (or any chatbot-mediated learning, for that matter) ultimately boils down to learning in a distinctive learning environment involving a human-machine interface. It, therefore, stands to reason that studies exploring the role of ChatGPT in language learning should connect with Instructed Second Language Acquisition (ISLA) research (Ellis, 1991; Loewen, 2020), an established resource on classroom learning of additional languages. Doing so would have a chain of advantages. First and foremost, it would eschew the need to reinvent the wheel, thereby elevating the researcher's point of departure at the outset. Second, it would open up methodological options. Most importantly, the resulting work can integrate with existing ISLA research to enhance our understanding of conditions conducive to language learning, both across and within learning contexts (e.g., human-human interaction, human-machine interaction).

Despite the ongoing debate in society writ large between “accelerators” and “doomers” over whether or not A.I. should or should not be allowed to move faster (Roose, 2023), the fact of the matter is that ChatGPT is already ubiquitous, gathering steam among language learners and instructors. This fact alone compels the field of SLA to engage in serious empirical inquiries into its functionalities and how learners and instructors use them as part of second language learning processes and outcomes.

Acknowledgments. Shaohua Fang contributed to the bibliographic search for this editorial. The Studies in Second Language Acquisition (SSLA) team of editors provided valuable feedback on a previous draft. Any errors or inaccuracies are solely my responsibility.

References

- Baskara, F., & Mukarto, F. (2023). Exploring the implications of ChatGPT for language learning in higher education. *Indonesian Journal of English Language Teaching and Applied Linguistics*, 7(2), 343–358. <https://files.eric.ed.gov/fulltext/EJ1391490.pdf>
- Bonner, E., Lege, R., & Frazier, E. (2023). Large language model-based artificial intelligence in the language classroom: Practical ideas for teaching. *Teaching English with Technology*, 23(1), 23–41. <https://eric.ed.gov/?id=EJ1383526>
- Ellis, R. (1991). *Second language acquisition and language pedagogy*. Multilingual Matters.
- Ellis, N. C., & Wulff, S. (2020). Usage-based approaches to L2 acquisition. In B. VanPatten, G. Keating, & S. Wulff (Eds.), *Theories in second language acquisition* (pp. 63–82). Routledge. <https://doi.org/10.4324/9780429503986-4>
- Fryer, L., Coniam, D., Carpenter, R., & L˘apușneanu, D. (2020). Bots for language learning now: Current and future directions. *Language Learning and Technology*, 24(2), 8–22. <https://www.lltjournal.org/item/10125-44719/>
- Gass, S., & Mackey, A. (2020). Input, interaction, and output in L2 acquisition. In B. VanPatten, G. Keating, & S. Wulff (Eds.), *Theories in second language acquisition* (pp. 192–222). Routledge. <https://doi.org/10.4324/9780429503986-9>
- Gregg, K. (1996). The logical and developmental problems of second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 49–81). Academic Press. <https://www.semanticscholar.org/paper/The-Logical-and-Developmental-Problems-of-Second-Gregg/bc27874b8a0860795c5db127d3b22044407eda99>
- Han, Z.-H., Kang, E., & Sok, S. (2023). The complexity epistemology and ontology in second language acquisition: A critical review. *Studies in Second Language Acquisition*, 45, 1388–1412. <https://doi.org/10.1017/S0272263122000420>
- Hong, W. (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation*, 5(1). Available online: <https://jeti.thewsu.org/index.php/cieti/article/view/103> (accessed on 10 August 2023).

- Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38, 237–257. <https://doi.org/10.1111/jcal.12610>
- Hulstijn, J. H. (2015). *Language proficiency in native and non-native speakers: Theory and research*. John Benjamins.
- Ji, H., Han, L., & Ko, Y. (2023). A systematic review of conversational AI in language education: Focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*, 55(1), 48–63. <https://doi.org/10.1080/15391523.2022.2142873>
- Kohnke, L., Moorhouse, B. I., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2). <https://doi.org/10.1177/00336882231162868>
- Lantolf, J., Poehner, M., & Thorne, S. (2020). Sociocultural theory and L2 development. In B. VanPatten, G. Keating, & S. Wulff (Eds.), *Theories in second language acquisition* (pp. 207–226). Routledge. https://www.researchgate.net/publication/313795407_Lantolf_J_Thorne_S_L_Poehner_M_2015_Sociocultural_Theory_and_Second_Language_Development_In_B_van_Patten_J_Williams_Eds_Theories_in_Second_Language_Acquisition_pp_207-226_New_York_Routledge
- Larsen-Freeman, L. (2020). Complex dynamic systems theory. In B. VanPatten, G. Keating, & S. Wulff (Eds.), *Theories in second language acquisition* (pp. 248–270). New York: Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9780429503986-11/complex-dynamic-systems-theory-diane-larsen-freeman>
- Li, B., Kou, X., & Bonk, C. (2023). Embracing the disrupted language teaching and learning field: Analyzing YouTube content creation related to ChatGPT. *Languages*, 8(197). <https://doi.org/10.3390/languages8030197>
- Loewen, S. (2020). Introduction to instructed second language acquisition (2nd ed.). Routledge.
- Metz, C. (2024). Robots learn, chatbots visualize: How 2024 will be A.I.'s 'leap forward.' *The New York Times*. Retrieved from <https://www.nytimes.com/2024/01/08/technology/ai-robots-chatbots-2024.html?smid=nytcore-ios-share&referringSource=articleShare>
- Roose, K. (2023). Many details of Sam Altman's ouster are murky. But some things are clear. *The New York Times*. Retrieved from <https://www.nytimes.com/2023/11/18/technology/sam-altman-open-ai.html>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?. *Journal of Applied Learning and Teaching*, 6(1), 342–363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Satariano, A., & Kang, C. (2023). How nations are losing a global race to tackle A.I.'s harms. *The New York Times*. Retrieved from <https://www.nytimes.com/2023/12/06/technology/ai-regulation-policies.html?searchResultPosition=3>
- The Southeastern Conference. (2023). When AI in higher education means more. *The Chronicle of Higher Education*. Retrieved from <https://sponsored.chronicle.com/when-ai-in-higher-education-means-more/index.html>
- Tseng, W., & Warschauer, M. (2023). AI-writing tools in education: if you can't beat them, join them. *Journal of China Computer-Assisted Language Learning*, 3(2), 258–262. <https://www.semanticscholar.org/paper/AI-writing-tools-in-education%3A-if-you-can-t-beat-Tseng-Warschauer/47ec3fc0e597f270d9b33dd08d78c1d8de1eb459>
- Van Geert, P. (2023). Some thoughts on dynamic systems modeling of L2 learning Front. *Phys*, 11. <https://doi.org/10.3389/fphy.2023.1186136>
- VanPatten, B., Keating, G., & Wulff, S. (Eds.). (2020). *Theories in second language acquisition*. Routledge.
- Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for second language writers of English as a second or foreign language. *Journal of Second Language Writing*, 62. <https://doi.org/10.1016/j.jslw.2023.101071>
- Xiao, F., Zhao, P., Sha, H., Yang, D., & Warshauer, M. (2023). Conversational agents in language learning. *Journal of China Computer-Assisted Language Learning*. <https://doi.org/10.1515/jccall-2022-0032>
- Zou, B., Reinders, H., Thomas, M., & Barr, D. (2023). Editorial: Using artificial intelligence technology for language learning. *Frontiers in Psychology*, 14(1287667). <https://doi.org/10.3389/fpsyg.2023.1287667>

Cite this article: Han, Z. (2024). Chatgpt in and for second language acquisition: a call for systematic research. *Studies in Second Language Acquisition*, 1–6. <https://doi.org/10.1017/S0272263124000111>