An Evaluation of Khanmigo, a Generative AI Tool, as a Computer-Assisted Language Learning App

Shamini Shetye
Teachers College, Columbia University

INTRODUCTION

Generative AI (GenAI) has recently gained great momentum and has been considered a potential educational resource by many learners. While there exists a plethora of commercially available Artificial Intelligence (AI) apps that serve as language learning tools, they are mostly developed without any adherence to Second Language Acquisition (SLA) theoretical underpinnings, and research supporting their potential for language learning mainly is based on learners' perceptions of AI tools (Han, 2024). There is also converging evidence that these tools do not provide an authentic language learning experience due to factors such as the unnaturalness of computer-generated voices used for interactions, the frequency of communication breakdowns, and the scripted nature of conversations, leading to ineffective communication or off-topic conversations that include meaningless sentences (Fryer & Carpenter, 2006; Huang et al., 2022). The attention then naturally diverts to GenAI due to its growing popularity. However, the potential of Gen AI as a language learning tool remains largely unexplored. With due consideration, this paper endeavors to gauge the potential of a recently launched Gen AI tool for its language learning potential, particularly "Khanmigo," launched in 2022 by the founders of Khan Academy (Anand, 2023). One of the reasons for evaluating this app is that Khan Academy (KhanAcademy.com), a non-profit organization, has been at the forefront of providing free quality education and ubiquitous access to online resources to interested learners of all ages and from diverse backgrounds. Anyone worldwide can learn at their own pace to enhance their knowledge across a wide variety of subjects, ranging from natural sciences to humanities.

Their recent introduction of Khanmigo, a GenAI-based educational app offered as part of a paid subscription with Khan Academy, is designed to help users accomplish a wide variety of tasks. As a tutor, it has been developed to assist or support learners through scaffolded learning activities. Even though the developers do not claim that Khanmigo is designed for language learning, its capability as a language learning tool would still merit further exploration based on research-based grounds. One of the rationales is that GenAI-powered tools, such as ChatGPT (Generative Pretrained Transformer) developed by Open AI, have been drawing attention to language learning due to its natural conversation generating and language modeling abilities, albeit the notoriety of ChatGPT in fabricating false and biased information (Bender et al., 2021; Han, 2024). However, Khanmigo, which is powered by ChatGPT's most sophisticated Large Language Model (LLM), GPT-4, via an Application Programming Interface (API), claims its superiority when compared to other AI tools because its content is developed by teachers. For example, it is the only tool that is built into the Khan Academy content library, which covers

topics like math, humanities, coding, sciences, and many more; and it is trained not to give away answers like other LLM-based AI tools do, but rather to guide learners to find answers as a teacher would.

Despite the technical sophistication in developing Khanmigo using GPT-4 model, the app needs to be evaluated like other previous language learning technologies for its appropriateness for certain contexts and groups of language learners. Moreover, a thorough review of any language learning tool must account for both context and learners since, as Chapelle (2001) notes, "an evaluation has to result in an argument indicating in what ways a particular CALL task is appropriate for particular learners at a given time" (p. 53).

This paper, therefore, provides an overview of Khanmigo's capabilities as a language learning tool using Chapelle's (2001) Evaluation Framework for CALL task appropriateness. First, the paper will briefly explain the evaluation methodology adopted. Next, it will describe the activities, target skills, and app interface. The app will then be evaluated using Chapelle's (2001) Framework, followed by a conclusion on the potential of Khanmigo as a language learning tool, specifically for learning French.

EVALUATION METHODOLOGY

Data Collection Procedures

The app was qualitatively analyzed by the researcher and author based on her interactions with the app for 17.5 hours using Chapelle's (2001) evaluation criteria for CALL task appropriateness, discussed below. The author examined the app and used it for her personal learning goals to advance her knowledge of French. The author could be considered a highly motivated high-beginner learner of French as a Second Language (L2) (e.g., A2 level per the Common European Framework of Reference; Council of Europe, 2001). The author is also a graduate student pursuing a doctoral degree in Applied Linguistics in the United States. Additionally, the author has a master's degree in Teaching English to Speakers of Other Languages/Applied Linguistics.

## Type of Activities, Targeted Skills, and Interface of Khanmigo

Khanmigo encompasses several different activities for learners to choose from. They vary based on the subject or topic offered or the nature of the task itself. The main categories of activities are referred to as follows: *Tutor Me, Refresh, Write, Debate, Chat, Play*. The *Tutor Me* format, as illustrated in Figure 1, is primarily designed to help learners with problem solving in math, science, or the humanities. Learners have to enter their questions, and Khanmigo will break the problem into its components and provide a series of steps to get the answer. It also offers more practice problems as needed.

FIGURE 1
The *Tutor Me* Interface of Khanmigo
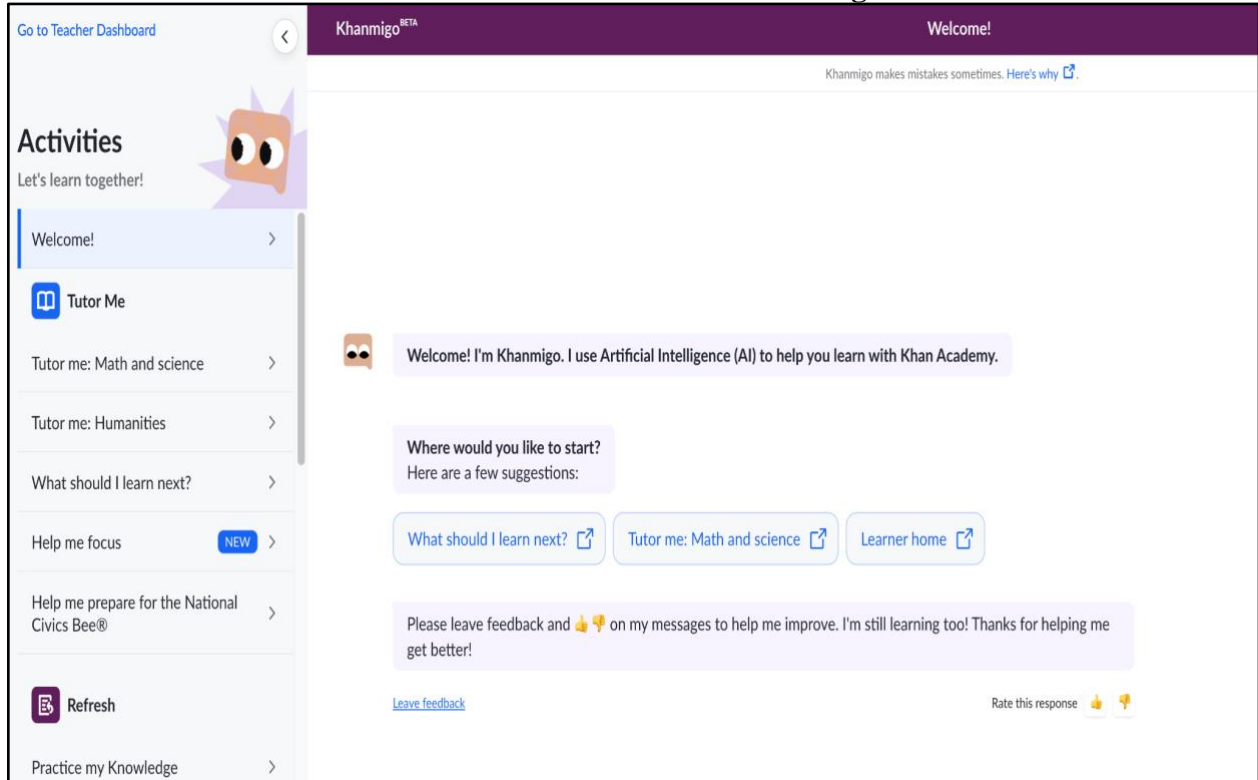
The *Refresh* format, as illustrated in Figure 2, allows learners to test their understanding of what they learned by taking a quiz. Khanmigo asks learners to enter their topic of interest and their grade level. Several questions will be posed, but one at a time. Feedback will be provided for learners' input, and more questions will be generated based on learners' previously demonstrated understanding of the subject.

FIGURE 2
**The *Refresh* Format of Khanmigo**
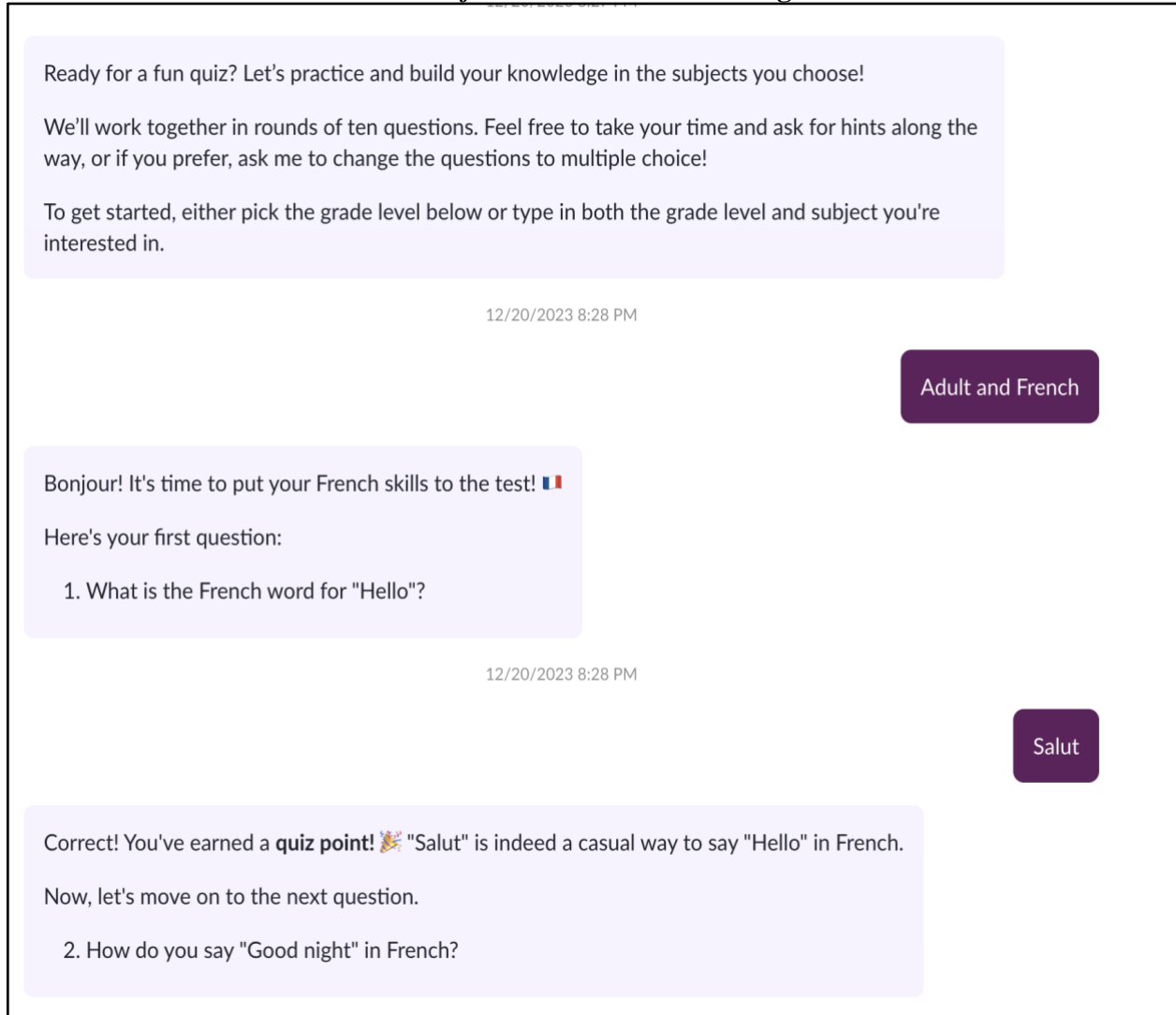
The *Write* format, as shown in Figure 3, deals with exposition writing. In this format, learners get feedback on their essays, or they are able to craft stories by brainstorming with Khanmigo. Besides general essay writing, this format specializes in brainstorming and providing feedback on users' admission essays.

The *Debate* format is specifically designed with young learners in mind, especially those in elementary, middle, and high school. Learners can brainstorm on different topics they provide or choose a topic from the selection provided by Khanmigo. The *Chat* format allows learners to choose a literary character or historical figure and have a simulated conversation with them. There are 29 fictional (e.g., Helen of Troy, Winnie the Pooh, Zeus) and 68 historical characters (e.g., Nikolas Tesla, Mahatma Gandhi, Plato, Napoleon Bonaparte) to choose from. The *Play* format offers learners the opportunity to play a variety of games involving the creation of new words or exploration of existing ones. The *Extra* section allows learners to converse with Khanmigo on various topics that spark their curiosity. They can navigate college admissions, discuss their personal interests, or seek academic or career-related advice. While each activity

has a specialized focus, the *Tutor Me, Write, and Extra* are more suited for learning and writing through multiple discussions or text construction with Khanmigo; therefore, these will be explored further in the study.

FIGURE 3
The *Write* Feature of Khanmigo
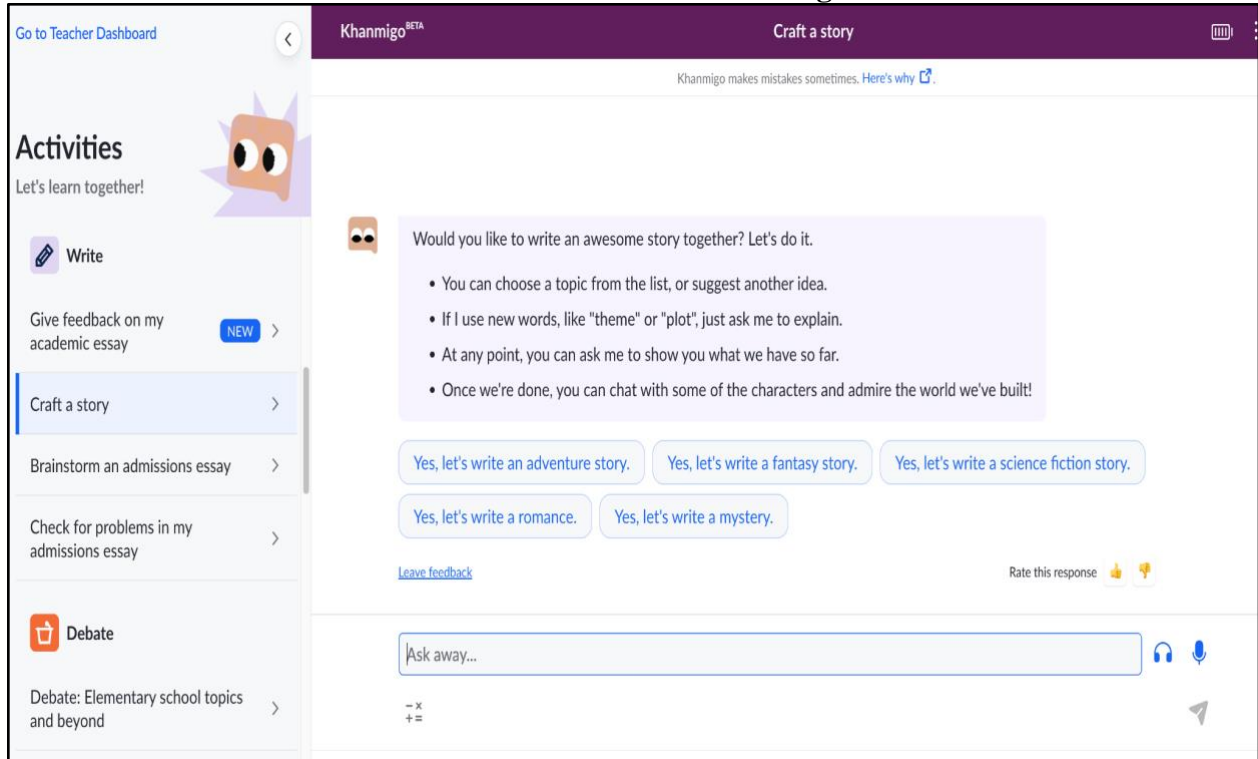
This app is also intended to serve as an autonomous learning tool and was not developed for language learning per se. It is not known which languages can be learned using the app. However, it immediately gets on task when asked if it can teach French, as depicted in Figure 4. Even though the app has the capability of speech-to-text and text-to-speech recognition, it is not available for learning French at the current time. It is more suitable for developing writing and reading skills. Despite this, learners can still ask questions about the pronunciation of words, and it will be spelled out, as illustrated in Figure 5.

Khanmigo's interface is easy to navigate and intuitive. The activities are all aligned on the left navigation pane, and the space in the center enables users to interact with the app depending on the activities they choose. Teachers and learners can both use the app based on the preferences specified. However, this evaluation will be presented from a learner's perspective. The app is available on multiple platforms, iOS and Android, and can be accessed via mobile devices.

FIGURE 4
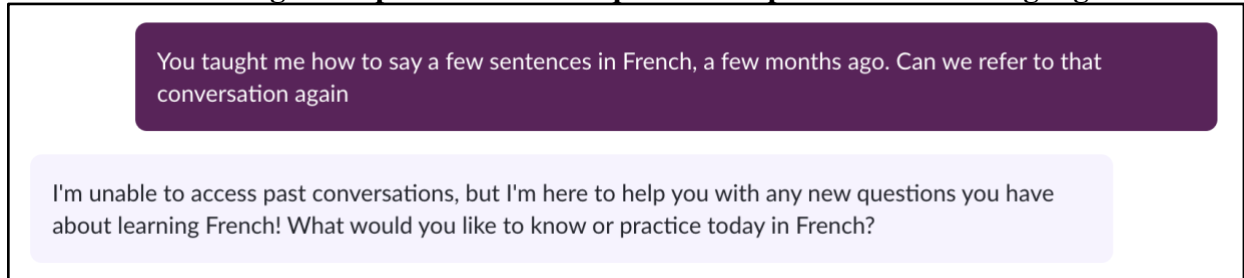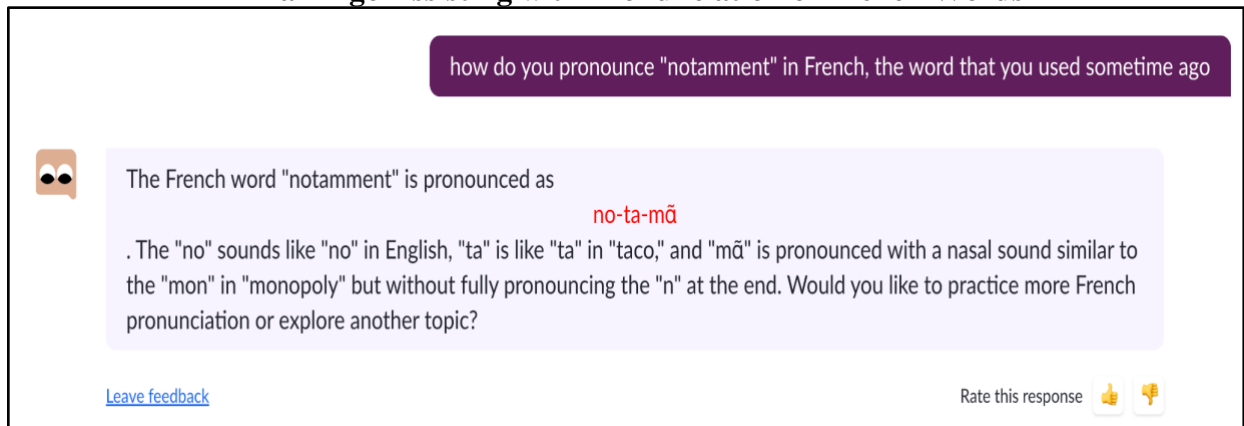**Khanmigo Accepts Learner's Request to Help with French Language**

FIGURE 5
Khanmigo Assisting with Pronunciation of French Words

EVALUATION OF KHANMIGO USING CHAPELLE'S (2001) FRAMEWORK

Chapelle's (2001) Framework delineates six criteria for assessing the task appropriateness of CALL Tools: Language Learning Potential, Learner Fit, Meaning Focus, Authenticity, Positive Impact, and Practicality. This framework is pertinent for analyzing CALL tools because it considers some theoretical tenets of Second Language Acquisition (SLA), particularly Instructed SLA (ISLA) and Applied Linguistics. SLA-based theoretical perspectives are considered useful for making evaluation arguments because of the ideal conditions they hypothesize for L2 learning through the use of technology (Chapelle, 2003). These theory-driven expectations can then guide evaluators to make relevant observations about the learning processes prompted by a given CALL tool, the learning outcomes yielded from this CALL tool or both. Since the findings are based on the theoretical hypotheses, the interpretations are considered meaningful, and they can be useful to support claims about the learning conditions, particularly the success of conditions (Chapelle, 2017).

Additionally, Chapelle's (2001) framework also considers other important factors, such as learners' individual differences, and integrates pedagogical perspectives that may be overlooked in other SLA theories (e.g., Intercultural Competence). However, the point is that if the evaluator specifies the pedagogical goals, the framework allows for the inclusion of such pedagogical perspectives (Chapelle, 2017). Moreover, using this framework, makes it possible to

do the evaluation on multiple levels: a judgmental level (i.e., to analyze if the software allows enough interactional opportunities), and the empirical level (i.e., by gathering data if the learner had enough opportunities to interact with the software).

While Chapelle's (2001) evaluation framework can be applied to all CALL tools, this paper contributes to the scholarly literature in this realm by applying the framework to gauge the potential of a particularly novel GenAI tool, Khanmigo, for language learning purposes through the six criteria in the framework. A summary of the six criteria in Chapelle's (2001) framework, the researcher's judgments on Khanmigo's potential as a language-learning app with respect to each criterion, and the evidence for each judgment is presented in Table 1. The findings are then discussed in further detail in the sections below.

TABLE 1
Summary of Khanmigo's Evaluation Using Chapelle's (2001) Framework

Criteria	Judgment	Evidence
Language Learning Potential *Degree of opportunity for beneficial focus on form*	Fully Supported	The interactions with Khanmigo provide opportunities to engage with abundant language input, produce meaningful output, receive adaptive feedback, and negotiate the meaning of language.
Learner Fit *Engagement opportunity with language considering learner characteristics*	Not Supported	The language used by Khanmigo and the topics selected by Khanmigo for discussions may be too advanced for beginner-level learners.
Meaning focus *Degree of learner's attention geared to the meaning of the language*	Partially Supported	Khanmigo offers learners Opportunities for two-way interactions focusing on meaning. However, the cognitive complexity of these interactions can be problematic for beginner-level learners.
Authenticity *Degree to which CALL tasks relate to those learners encounter outside the instructional setting.*	Fully Supported	Nearly all of the activities offered in Khanmigo, except for *Chat,* have the potential to engage learners in authentic activities.
Positive Impact *Positive effects of the CALL activity on participants*	Not Supported	Khanmigo has the potential to help learners develop their metacognitive strategies tacitly. However, these activities in Khanmigo lack a clear pragmatic focus.

Practicality	Partially Supported	The associated Cost ($4 per month, as of May 2024) is fairly affordable. Additionally, Khanmigo has a user-friendly interface, is portable, and is easy to install. It is also possible to add teachers as supervisors to monitor learner progress.
Adequacy of resources to support the CALL activities		

Language Learning Potential

This evaluation criterion concerns the degree to which the task can involve learners with a beneficial focus on form (Chapelle, 2001, p. 55). Khanmigo provides ample opportunities for learners to engage in communicative tasks that focus on meaning. All task types encourage learners to pick up a topic of interest and communicate without explicit focus on forms or grammatical corrections. The feedback provided by Khanmigo mostly focuses on verifying the meaning and the context. Learners can also ask specific questions to get feedback on their form.

For instance, in one subtask categorized under the *Extra* format, referred to as *Ignite My Curiosity*, the learner can choose one of the questions Khanmigo provides to get more information on the topic. For example, the researcher requested information in French and was provided with a list of topics in the form of questions. The researcher selected the following question, "Comment le français a-t-il évolué au fil du temps?" (English translation: How has the French evolved over time?). Khanmigo responded by providing a brief history of the language's evolution over time, as illustrated in Figure 6, and with an English translation, provided in Figure 7. Although the learner was unfamiliar with a few words (6 words), when asked for explanations in English or in French, Khanmigo provided responses adding adequate contexts without an explicit focus on forms, as illustrated in Figure 8. Interactions with the app provided meaningful ways for the learner to engage in enriching conversations that not only developed learners' knowledge of the language but also provided contextual information about the topic.

From the interaction-approach perspective (Gass & Mackey, 2020), which subsumes multiple influential hypotheses in SLA, such as the Input Hypothesis (Krashen, 1985), Output Hypothesis (Swain, 1995), and Interaction Hypothesis (Long, 1983), negotiation is an essential element of interaction as it draws learners' attention to input, output, and feedback, which all promote language acquisition. Khanmigo affords numerous opportunities for learners to engage in negotiation through abundant input, opportunities for extended output production in contextualized interactions coupled with the ability for learners to ask unlimited questions and get real-time feedback for all of their responses. Furthermore, the feedback is adapted based on learners' previous knowledge or responses. Therefore, considering all these aspects, the author has concluded that the evidence from the app fully supports the language learning potential criteria.

FIGURE 6
Khanmigo's Demonstrating Beneficial Focus on Form

Comment le français a-t-il évolué au fil du temps?

Ah, l'évolution du français, quelle histoire fascinante! 📖🇫🇷

Le français que nous connaissons aujourd'hui a des racines profondes, remontant au latin parlé dans l'Empire romain. Au fil des siècles, le latin a évolué en plusieurs langues différentes, dont le français.

Au Moyen Âge, le français était très différent de ce qu'il est aujourd'hui. Il y avait de nombreuses variantes régionales et beaucoup de variations dans la grammaire et le vocabulaire. C'est à cette époque que le français a commencé à influencer l'anglais, après l'invasion normande de l'Angleterre en 1066.

À l'époque moderne, le français a continué à évoluer, avec l'introduction de nouveaux mots et expressions pour refléter les changements dans la société et la technologie.

FIGURE 7
Translation of Text in Figure 6 from French to English

How has French evolved over time?

Ah, the evolution of French, what a fascinating story! 📖🇫🇷

The French we know today has deep roots, going back to the Latin spoken in the Roman Empire. Over the centuries, Latin evolved into several different languages, including French.

In the Middle Ages, French was very different from what it is today. There were many regional variants and a lot of variations in grammar and vocabulary. It was during this time that French began to influence English, after the Norman invasion of England in 1066.

In the modern era, French continued to evolve, with the introduction of new words and expressions to reflect changes in society and technology.

Would you like to explore more about the history of the French language, or is there another topic you're curious about?

FIGURE 8
Learner Confirming Meaning of Unfamiliar Words
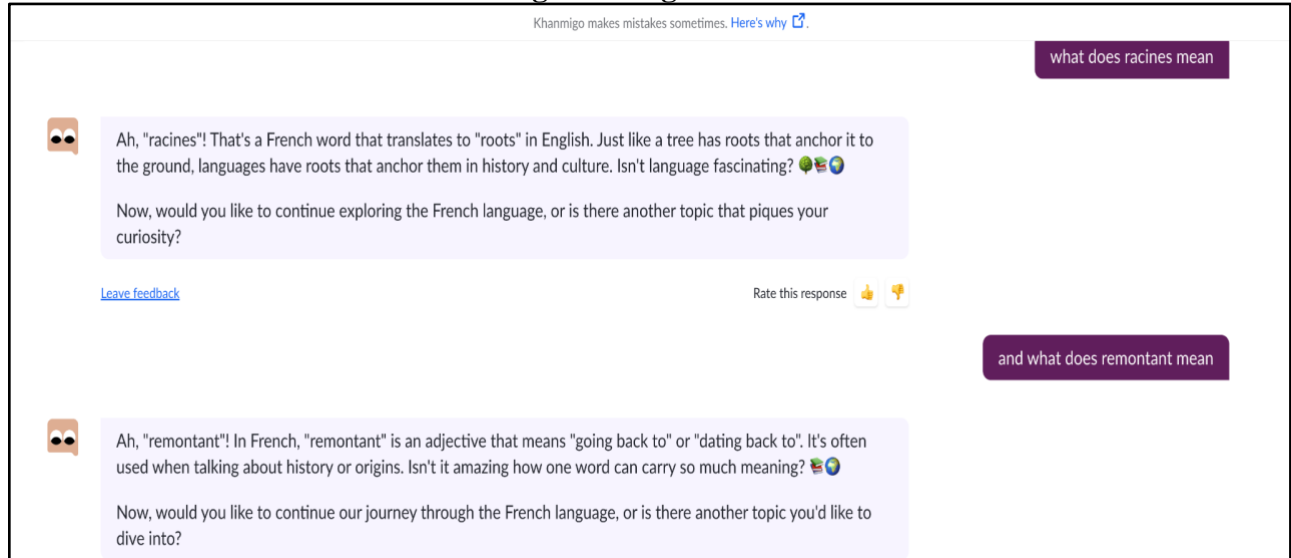
Learner Fit

This criterion considers individual learner variations and learners' linguistic ability levels as well as certain non-linguistic characteristics (Chapelle, 2001, p. 55). The tasks should be appropriate to learners' levels and should not be exceedingly challenging or extremely familiar to learners. Other characteristics such as learning style, age, and willingness to communicate also ought to be considered. Khanmigo may pose a challenge here when it comes to matching learner fit. Learners at the beginner level may feel overwhelmed with the amount of novel language if the input provided by the app is at an advanced level and is, therefore, incomprehensible to them.

Per Pienemann's (1998) Processability Theory, learners can only comprehend and produce linguistic forms that the language processor can manage at their current stage of development (i.e., there are universal developmental sequences, and learners follow a staged development trajectory). So, even if learners are motivated enough to learn, the advanced input may not benefit L2 development. Learners may find the task daunting if Khanmigo chooses a topic that may not interest them or is unfamiliar.

Moreover, research in usage-based approaches has shown that increased salience in input in terms of word familiarity, meaningfulness, and word concreteness facilitates language acquisition (see Crossley et al., 2016). Khanmigo's potential to interact with users in a language that has more salience is debatable, particularly in the case of beginner-level learners. Unless they are highly motivated and understand that they need to probe the app with appropriate queries to match their level and interests, learners may withdraw their attention and not stay focused on learning. Moreover, in unsupervised or untutored learning conditions, if adequate support is not available to learners, it may further hinder their progress. For instance, if we look at the conversation in Figures 6 and 7 above, which focus on meaning, and focus on a topic that may be advanced for some learners, it may be challenging for a beginner learner of French to completely understand the meaning of the passage, unless they ask further questions about the meaning of certain words, as the researcher did, as shown in Figure 8 above. That may require more motivation and willingness to pursue the communication further. Additionally, although

Khanmigo provides options for learners to choose a topic that interests them, if the conversation continues in an area or topic that is unfamiliar to them, it may become cognitively demanding, which may further disrupt the chance of learning further. Thus, under unsupervised or untutored conditions, the app may not benefit learners of all ages or L2 proficiency levels. Therefore, the criterion of learner fit has not been met and is not supported by the evidence reported about this app.

Meaning Focus

This criterion relates to the extent to which the task or activity is meaning-focused so that it encourages the learner to focus on form. To evaluate if the tasks are meaning-focused, it would be worthwhile to follow the guidelines that Pica et al. (1993) delineate for making such determinations. First, the task should address some information gap exchange. Ideally, it should be a two-way task. Second, the outcome should be closed, i.e., there is not more than one possible outcome. Third, the task should not be familiar to the learner. Fourth, the topics must be humane or ethical, i.e., they are appropriate for learners and not culturally insensitive. Lastly, the cognitive complexity must be low, and the discourse domain must be narrative, which can allow for more meaning-oriented tasks.

If we bear in mind guidelines from Pica et al. (1993), Khanmigo does include tasks with information gaps, which are two-way and often encourage learners to interact. The outcome may or may not be closed and would depend on the learners and their motivation to complete the task. In terms of topic familiarity, narrative discourse, and their ethical nature, it would again allow learners to choose topics that they would be interested in. Cognitive complexity can be high if learners do not generate a suitable prompt. For example, as depicted in Figure 9, followed by an English translation in Figure 10, although the learner chose the topic of their interest, the text generated may be too cognitively demanding if the learner is not familiar with the subject.

Nevertheless, if these guidelines are taken as a heuristic, Khanmigo has tremendous potential to engage its interlocutors in meaning-focus tasks, provided if the cognitive complexity is matched to learners' levels. This may depend on the ability of the AI-generation tool to adapt to learners' abilities. Consequently, the criteria for being meaning-focused has been met, to a certain extent, by Khanmigo, and is partially supported by the evidence.

FIGURE 9
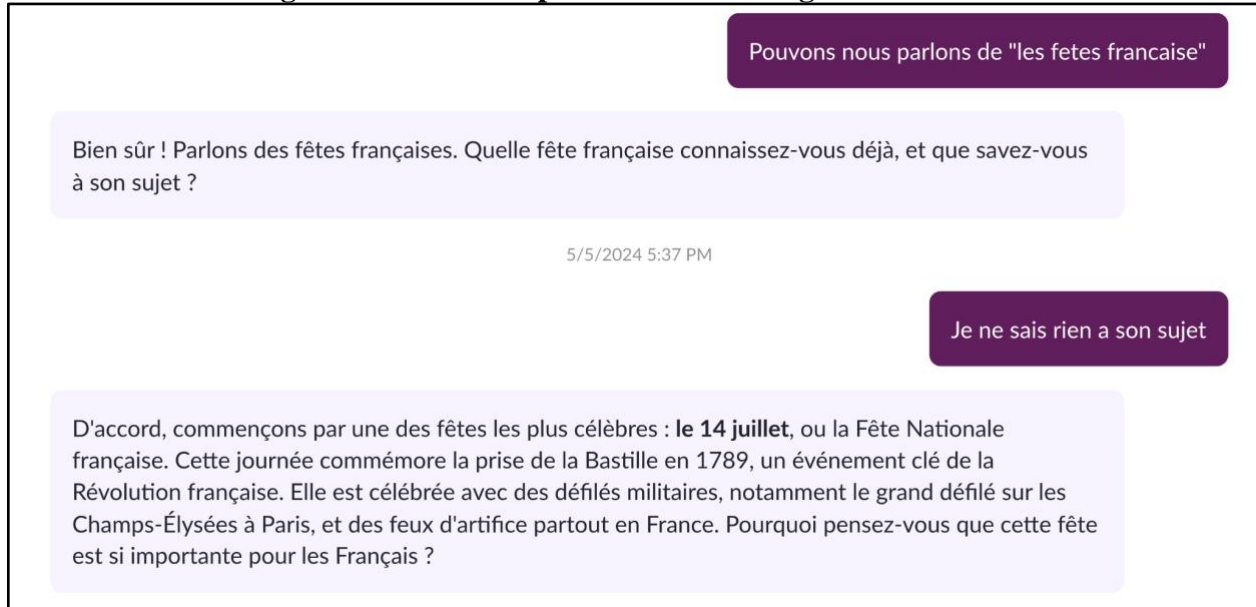**Khanmigo Introduces a Topic Not Fit for a Beginner-Level Learner**

FIGURE 10
Translation of Text in Figure 9 from French to English

Authenticity

This criterion refers to the degree to which the learning task mirrors scenarios in the real world which learners are more likely to experience outside the classroom setting. There is consensus among researchers about the possibility of authentic tasks to facilitate language development (Lightbown et al., 1993; Long, 2017; Savignon, 2018). Here, Khanmigo can be seen as an asset as it can generate texts that are more authentic in nature. Learners can choose a subject that they may need to practice. It could be humanities or even science subjects. Learners, while learning a language, can focus on specific purposes that meet their current needs. Moreover, Khanmigo claims to model a Socratic approach to teaching in that it does not provide students with answers, but guides their learning by asking thought-provoking, open-ended

questions, which then lead them to learn more. This pedagogical approach may resonate with many classroom teachers and speaks to the authenticity of the method.

Another consideration is the modality of communication. Although only writing and reading, not speaking or listening, were available for French learning in Khanmigo, the researcher does not see this as a demerit. Admittedly, it would be ideal to integrate all skills in learning activities, but that is not always practical to implement. Therefore, other than the activity that encourages learners to have simulated conversations with literary or historical figures, most activities resemble those encountered in real-life settings. Therefore, the researcher considers the criterion for authenticity has been met and is fully supported by the evidence.

Positive Impact

This criterion refers to the degree to which the tasks can provide learners with opportunities to hone their metacognitive strategies, or "techniques used to help students understand the way they learn and do this more effectively" (Hornby & Greaves, 2022, p.2), and gain pragmatic competence, which is "the knowledge of form-function-context mappings-which forms to use for what communicative functions in what social contexts" (Taguchi, 2019, p. 13). Khanmigo provides scaffolding for learners to complete a task. This, in a way, helps learners to develop their metacognitive skills, as it gives time for learners to reflect and add content. However, the method is tacit and is not explicitly taught to learners when they use the app. As for pragmatic competence, Khanmigo is not necessarily able to improve learners' skills in pragmatics, unless learners ask questions related to pragmatics. So, the evidence from Khanmigo does not support the positive impact criterion.

Practicality

When it comes to practicality, several facets merit attention. First is the concern about access to the hardware or software. Khanmigo is not an open-source app and comes with a paid subscription ($4 monthly as of May 2024). Consequently, not all learners and educators can afford to pay a monthly fee. Secondly, although Khanmigo is designed to be an autonomous learning tool, practicality also means that knowledgeable personnel need to be around to help with learning. This may not always be feasible. However, Khanmigo has a feature that allows educators to log in and review students' progress if the students are assigned to a specific class. The toggle mode on the app allows users to switch between two modes: learners or students and teacher. This still makes an option worth considering. As far as the usability of the software, it is considerably intuitive, and learners and educators can access the app with minimal difficulty. Concerning the portability of the app, it can be accessed from mobile platforms, Android and iOS devices. It can be installed on any device. So, when it comes to practicality, the evidence from Khanmigo partially supports this criterion.

CONCLUSION

The purpose of this paper was to evaluate Khanmigo, a novel GenAI-enabled language learning app, through the lens of Chapelle's (2001) criteria for evaluating CALL tools. After about 17.5 hours of interactions with Khanmigo to learn French, through various activities it

offers, the researcher is somewhat skeptical and feels mixed about the potential of Khanmigo as a robust language learning tool, particularly for beginner L2 French learners. The app may still benefit advanced language learners if they are able to learn a language through a self-guided process. While the *language learning potential* is immense through contextual interactions that focus on form, it may not be suitable for all learners. In other words, it does not support the criterion for *learner fit*. It may be more suitable for advanced learners who can converse on versatile topics and understand the advanced language used by the app. While there are reasonable opportunities for learners to engage in *meaning-focused* activities or *authentic* activities, and the *practicality* of the app seems high enough for it to be widely used, the lack of potential for learners to gain pragmatic skills is a concern that needs to be considered. Thus, the evidence from the app does not fully support the criterion of *positive impact*. While Khanmigo was not developed to be used as a language learning tool, it does show some promise, and developers should consider how they could add some of the features mentioned in this report to improve the app so that language learners can utilize it further.

However, there are some study limitations that need to be acknowledged. First, the researcher examined the app's potential as a language learning tool, particularly for French, even though the app was not specifically designed for the purpose. Evaluation of the app for learning English may yield varied results. The results may also vary with new software development to improve the app. Additionally, the researchers' reported impressions were based on a qualitative analysis vis-à-vis Chapelle's (2001) framework. To validate these reported interpretations, further empirical research, experimental and non-experimental, would be needed. Moreover, it would also be worth conducting systematic research examining long-term language learning gains through the longitudinal use of the AI tool.

ACKNOWLEDGEMENTS

REFERENCES

Anand, P. (2023, March 16). Khan Academy creates GPT-4 based helper Khanmigo marking formal entry of AI into education. *Dataquest.* https://www.dqindia.com/khan-academy-creates-gpt-4-based-helper-khanmigo-marking-formal-entry-of-ai-into-education/

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610–623). Association for Computing Machinery.

Chapelle, C. A. (2001). Principles for CALL evaluation (pp. 51-94). In *Computer applications in second language acquisition: Foundations for teaching, testing, and research.* Cambridge University Press.

Chapelle, C. A. (2003). *English language learning and technology*: *Lectures onaApplied linguistics in the age of information and communication technology*. John Benjamins Publishing.

Chapelle, C. A. (2017). Evaluation of technology and language learning. In C.A. Chapelle, & Sauro, S. (Eds.), *The handbook of technology and second language teaching and learning*. https://doi-org.ezproxy.cul.columbia.edu/10.1002/9781118914069.ch25

Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment.* Cambridge University Press.

Crossley, S., Kyle, K., & Salsbury, T. (2016). A usage-based investigation of L2 lexical acquisition: The role of input and output. *The Modern Language Journal*, *100*(3), 702–715. https://doi.org/10.1111/modl.12344

Fryer, L. & Carpenter, R. (2006). Bots as language learning tools. *Language Learning & Technology*, *10*, 8–14.

Gass, S., & Mackey, A. (2020). Input, interaction, and output in L2 acquisition. In B. VanPatten, Keating, G., & Wulff, S. (Eds.), *Theories of second language acquisition: An introduction*, 3rd ed. (pp. 192–222). Routledge.

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

Hornby, G. & Greaves, D. (2022). *Essential evidence based teaching strategies: Ensuring optimal academic achievement for students*. Springer.

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. *Computer Assisted Learning, 38*, 237–257. https://doi.org/10.1111/jcal.12610

Krashen, S. (1985). *The input hypothesis: Issues and implications.* Longman.

Lightbown, P. M., Spada, N., & White, L. (Eds.). (1993). *The role of instruction in second language acquisition*. Cambridge University Press.

Long, M. H. (1983). Native speaker/non-native speaker conversation in the second language classroom. In M. Clarke and J. Handscombe (Eds), *On TESOL '82: Pacific perspectives on language and teaching*. TESOL.

Long, M. H. (2017). *Problems in second language acquisition*. Routledge.

Pienemann, M. (1998). *Language processing and second language development*. Johns Benjamins.

Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language teaching and research. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 1–34). Multilingual Matters.

Savignon, S. (2018). Communicative competence. *The TESOL encyclopedia of English language teaching*, 1–7. https://doi.org/10.1002/9781118784235.eelt0047

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H.G. Widdowson* (pp. 125–144). Oxford University Press.

Taguchi, N. (2019). Second language acquisition and pragmatics: An overview. In N. Taguchi (Ed.), *The Routledge handbook of second language acquisition and pragmatics* (pp. 1-14). Routledge.

Shamini Shetye is an Ed.D. student in Applied Linguistics/TESOL and a doctoral fellow in the Community Language Program at Teachers College, Columbia University. Her research interests include viewing L2 lexical development through Complex Dynamic Systems Theory (CDST) and technology-mediated language learning. She is an active part of the AL TESOL Language & Technology Research Group at Teachers College. Correspondence should be sent to Shamini Shetye, Email: sus2004@tc.columbia.edu.