

This work is licensed
under a Creative
Commons Attribution
4.0 International
License.

Evaluating Synthesia as a CALL tool in university-level English for academic purposes course: A case study in Hong Kong

 Frankie Har

The Hong Kong Polytechnic University, Hong Kong
frankie.tk.har@polyu.edu.hk

Along with other AI-powered video generation platforms, Synthesia is a new Computer-Assisted Language Learning (CALL) resource which enables teachers to create interactive learning experiences for their students. The study investigates Synthesia technology applications for EAP teaching at a Hong Kong government-funded university to determine its educational advantages and technical challenges. Prior research about AI applications in language education exists but studies on how EAP students use avatar-based video tools are limited. The research investigates how Synthesia functions as an academic English learning tool by gathering data from students and their teachers. Qualitative data in the research shows students like Synthesia's user-friendly interface and ability to customize content, but students also express doubts about its authenticity, cognitive load, and assessment alignment. The research findings carry two key pedagogical implications: the need to integrate AI literacy into EAP curricula and the importance of aligning CALL tools with course learning outcomes to ensure meaningful and ethical use.

Keywords: AI-powered video generation platforms, cognitive load, Computer Assisted Language Learning (CALL), EAP, Synthesia

Introduction

Artificial Intelligence (AI) supports English for Academic Purposes (EAP) courses at universities because language education institutions now widely employ this technology. Among emerging AI tools, Synthesia, which generates AI videos through avatars, has become popular because it shows promise for improving multimodal learning experiences. This research study examines the pedagogical affordances and constraints of using Synthesia as a Computer-Assisted Language Learning (CALL) tool at a government-funded university in Hong Kong. The research takes place in

the digital transformation environment of higher education where instructors need to merge new technology with established teaching methods.

The research investigates how Synthesia technology enhanced EAP students' academic communication abilities through its application in presentation and role-play activities. It aims to evaluate the tool's effectiveness in fostering student engagement, reducing cognitive load, and promoting active learning, which are the three key principles identified by Brame (2017) for effective video-based instruction. The research also examines how instructors face time constraints and technical difficulties which affect their ability to create and distribute AI-based learning resources.

Literature review

The development of Computer-Assisted Language Learning (CALL) has moved from basic practice drill software to advanced platforms that support learning through multimodal and interactive approaches. Recent advances in Generative AI have further accelerated this shift by enabling on-demand video-based instructional content. The AI video generator Synthesia operates in line with Mayer's (2020) multimedia learning theory and Paivio's (1990) dual coding theory, both of which suggest that combining visual and auditory input enhances comprehension and memory retention. Empirical studies have shown that video-based instruction can support learners' listening and reading skills, as well as grammar and vocabulary development (Zhang & Zou, 2022). In addition, learners are better able to manage cognitive load when studying independently through videos that follow established instructional design principles (Castro-Alonso & Sweller, 2021).

Recent research on AI-powered video platforms for English language teaching further supports the pedagogical potential of avatar-based video tools when they are grounded in theory and integrated into well-designed learning tasks. For example, Har and Javier's (2026) critical review of the AI video platform HeyGen highlights how AI-generated videos can provide multimodal input, scaffold language learning, and increase learner engagement when aligned with established Second Language Acquisition (SLA) principles. Their analysis suggests that such platforms are most effective when used to support instructional delivery and guided practice rather than as replacements for human-mediated interaction.

Despite these advantages, AI-powered tools do not automatically guarantee pedagogical effectiveness. Brame (2017) argues that successful educational video design depends on three core elements: effective cognitive load management, sustained learner engagement, and opportunities for active learning. In a similar vein, Har and Javier (2026) caution that avatar-based video tools require careful pedagogical alignment, as over-reliance on AI-generated delivery may limit opportunities for communicative practice and higher-order academic language use. Within the Hong Kong EAP context, this highlights the need to critically evaluate the use of Synthesia in relation to academic literacy and critical thinking outcomes that form the foundation of the curriculum. Recent research in Hong Kong on inclusive CALL practices has further shown how digital tools can support learners from diverse backgrounds, including those from low-income families (Shi & Cheung, 2025). However, few empirical

studies have examined avatar-based video platforms such as Synthesia specifically in EAP settings, making the present case study a timely contribution to the field.

Methodology

This study employed a qualitative case study approach to examine the use of Synthesia as a Computer-Assisted Language Learning (CALL) tool. The course aimed to enhance students’ academic presentation skills through Synthesia video creation for demonstrations and simulation work. Focus group interviews were conducted to gather information from instructor (n=6) and student feedback (n=38) to understand their viewpoints. Thematic analysis (Braun & Clarke, 2006) was used in the research to detect recurring patterns in student and instructor answers about usability, engagement, and educational alignment. The following questions were asked:

- Q1) How easy or difficult was it for teachers to create and for students to interact with Synthesia videos compared to traditional video or presentation tools?
- Q2) Do Synthesia-generated videos hold the students’ attention more effectively than other formats (e.g., live lectures, slides, or recorded videos)? Why or why not?
- Q3) In what ways do Synthesia videos support or hinder the learning objectives in academic language education?

With respect to the thematic analysis, a total of 44 responses were coded with the aid of qualitative analysis software, NVivo 14, and 3 themes were identified. The thematic map showing responses is illustrated in Figure 1. Emerging from the interview data, the 3 themes offer significant insights into the ESL student learning and instructor teaching experiences with the support of Synthesia.

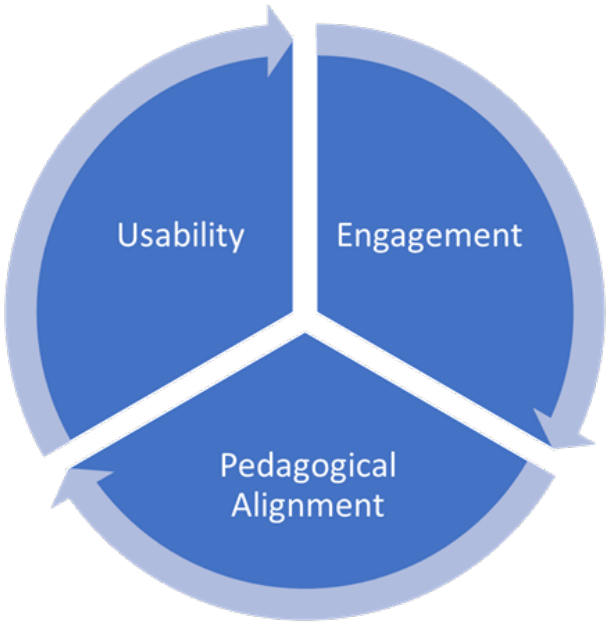


Figure 1. Thematic map of the use of Synthesia in EAP classes

Results and discussion

This section presents and discusses the findings derived from qualitative feedback provided by both students and instructors following their use of Synthesia in an English for Academic Purposes (EAP) context. The results are organized thematically to reflect key dimensions relevant to technology – enhanced language learning, namely usability, engagement, and pedagogical alignment with course objectives. Participants' comments are used to illustrate shared perceptions, perceived benefits, and emerging concerns, allowing for an in-depth examination of how Synthesia supports, and in some cases constrains, English language teaching and learning at the tertiary level.

Usability

From the following comments, one can infer the usability of the software:

Synthesia is really easy to use. I don't need any video editing skills to create presentations, and the avatars help me practise listening to different accents in English. (Student 1)

I like how I can type my script and get a video instantly. It saves time and lets me focus more on the content of my English lessons. (Teacher 1)

It is evident that the users of Synthesia express positive feedback about the platform because it works well for English language teaching and learning between teachers and students. The platform provides users with complete accessibility because they can create videos through script typing without needing video editing expertise. The platform allows teachers to concentrate on learners' content delivery because it simplifies the production process. The platform enables users to practice listening through avatar-based accent exposure which helps them develop their English language skills.

Engagement

The following sentences given by both instructors and students can infer the engagement made by Synthesia:

The variety of accents in Synthesia helps me get used to different English speakers, which is great for real-world communication. (Student 2)

Synthesia allows me to create engaging content quickly. Students respond well to the visual format, and it keeps them interested. (Teacher 2)

It's a great tool for flipped classrooms. I plan to give video assignments for homework while using classroom hours to help students develop their language skills at a more advanced level. (Teacher 3)

Based on the three comments, student engagement with Synthesia in English Language Learning is notably high. The learners benefit from different accents because it helps them develop their ability to recognize various English speakers

which they need for actual communication. The visual and interactive format of Synthesia videos helps students stay focused and interested while the tool allows teachers to use flipped classroom approaches for independent content study and better classroom activities. These insights reflect strong learner motivation and active participation driven by Synthesia's engaging features.

Pedagogical alignment

The following comments describe the degree of pedagogical alignment with the EAP course:

The customizable visually engaging content from Synthesia enables students to learn through multiple methods while they practice their academic communication skills. (Student 3)

While Synthesia is useful, the avatar-based delivery lacks authenticity, which may hinder the development of natural speech patterns in learners. (Student 4)

The implementation of Synthesia as a replacement for live interaction will fail to achieve its intended goals because it needs proper integration with instructor feedback to maintain alignment with essential learning objectives. (Teacher 4)

Based on the three comments provided, Synthesia shows both promise and limitations in its pedagogical alignment with an English for Academic Purposes (EAP) course. The instructors value Synthesia because it produces personalized visual materials which help students learn through multiple methods. The method supports EAP objectives which emphasize clarity, structure, and exposure to academic spoken discourse. The use of avatars for delivery, however, raises questions about their pedagogic alignment because an over-reliance on such software could prevent students from developing their speaking abilities. The instructors warn that Synthesia functions best as an addition to live classes because students need direct human interaction to achieve their EAP learning goals. The tool Synthesia supports EAP teaching methods when teachers use it as a supplementary tool, but they need to implement it correctly to prevent weaken essential communication abilities.

Conclusion and pedagogical implications

This study investigated the pedagogical potential of Synthesia within an English for Academic Purposes (EAP) setting, with particular attention to usability, learner engagement, and pedagogical alignment. Overall, the findings indicate that Synthesia can effectively support EAP instruction when its affordances are deliberately aligned with curricular goals and embedded within a comprehensive instructional design.

With respect to usability, both instructors and students consistently highlighted Synthesia's intuitive design and low technical threshold. The ability to generate videos by simply typing a script was perceived as particularly beneficial, as it removed the need for specialist video-editing skills and reduced preparation time for instructors. This ease of use allowed teachers to focus on content quality and instructional planning, while students benefitted from clear, structured academic input delivered

in a visually accessible format. These findings indicate that Synthesia is well suited for creating instructional materials in EAP courses, especially in contexts where instructors juggle multiple teaching responsibilities and require efficient tools to support material development.

In terms of learner engagement, the results demonstrate that Synthesia's visual format and exposure to varied English accents contributed positively to students' learning experiences. Students reported heightened interest and sustained attention when engaging with AI-generated videos, while instructors observed increased responsiveness to visually mediated content. Importantly, teachers also identified Synthesia as a useful tool for supporting flipped classroom approaches, whereby students engage with video-based input independently and reserve face-to-face class time for higher-level language practice. These findings suggest that Synthesia can enhance engagement by promoting active and flexible learning patterns, particularly when used to complement interactive classroom activities rather than to replace them.

However, as discussed in Section 4.3, the study also identifies important considerations regarding pedagogical alignment. While Synthesia was valued for its capacity to model academic language and present course-relevant content, both students and instructors cautioned against over-reliance on avatar-based delivery. Concerns centered on the extent to which passive consumption of AI-generated videos alone could support the development of learners' own spoken academic skills. The findings underline that Synthesia functions most effectively when embedded within a pedagogically coherent design that includes opportunities for live interaction, peer collaboration, and instructor feedback. In this sense, pedagogical alignment depends not on the tool itself, but on how instructors integrate it into task design and learning sequences that foreground communicative practice.

These findings lead to two key pedagogical implications. First, AI literacy should be embedded into EAP curricula from an early stage. Students need structured opportunities to evaluate the affordances and limitations of AI-mediated tools, including ethical considerations such as transparency, content ownership, and potential bias. Developing such critical awareness ensures that learners engage with AI technologies as informed users rather than passive consumers. Second, instructional use of CALL tools such as Synthesia must be guided by learning outcomes rather than technological novelty. In the context of this study, Synthesia was most effective when used to deliver instructional input and support engagement, while core academic speaking skills were developed through human-mediated interaction. Aligning task design, assessment criteria, and reflective activities with these principles is essential to ensuring that AI-supported tools meaningfully contribute to EAP learning objectives.

Overall, the study reinforces the view that Synthesia can appropriately be positioned as a supplementary instructional resource within a blended EAP pedagogy. When thoughtfully integrated, it can enhance usability and engagement while maintaining strong pedagogical alignment with the communicative goals of academic English instruction.

Acknowledgment

ChatGPT has been used for proofreading the language in this article. However, it should be noted that ChatGPT was not involved in generating any ideas. Any errors, if present, are solely those of the author.

References

- Brame, C. J. (2017). Effective educational videos: Principles and guidelines for maximizing student learning from video content. *CBE – Life Sciences Education*, 15(4), es6. <https://doi.org/10.1187/cbe.16-03-0125>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Castro-Alonso, J. C., & Sweller, J. (2021). The modality principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 261–267). Cambridge University Press. <https://doi.org/10.1017/9781108894333.026>
- Har, F. & Javier, D.R.C. (2026). HeyGen's AI video platform for English language teaching. (2026). *The Asian Journal of Applied Linguistics*, 10(1), 1313. <https://caes.hku.hk/ajal/index.php/ajal/article/view/1313>
- Mayer, R. E. (2020). *Multimedia learning* (3rd ed.). Cambridge University Press. <https://doi.org/10.1017/9781316941355>
- Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195066661.001.0001>
- Shi, H., & Cheung, L. M.E. (2025). Service-learning as inclusive CALL: A Hong Kong case study. *Computer-Assisted Language Learning Electronic Journal*, 26(2), 54–82. <https://callej.org/index.php/journal/article/download/571/426>
- Zhang, R., & Zou, D. (2022). Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning. *Computer Assisted Language Learning*, 35(4), 696–742. <https://doi.org/10.1080/09588221.2020.1744666>

Author biodata

Frankie Har has extensive experience in English language teaching, with research interests in AI and Gamification in ELT. He serves as a reviewer for leading journals including *Computers and Education*, *Computers and Education Open*, *Computers and Education: Artificial Intelligence*, *Computers and Education: X Reality*, *Computer Assisted Language Learning*, *RELC Journal*, *Innovation in Language Learning and Teaching*, and *Technology in Language Teaching & Learning*. His most recent and forthcoming publications include “Teachers’ Profiles of Inclusive Generative Artificial Intelligence Use in Language Education: An Ecological Mixed-Methods Study” (in press), “Navigating Professional Growth in AI in Language Education: Narratives from a University in Hong Kong” (in press), “Teaching English as a Foreign Language in the RenAIssance: An Exploration of Thai pre-service EFL teachers’ perceptions of AI in ELT Classroom” (in press), “ESL Pedagogy in Transition: Learners’ Perceptions of YouGlish for Pronunciation Focused Input” (in press), “ChatGPT 5.2 for multi-modal language teaching and learning: A two-year update” (2026), “HeyGen’s AI video

platform for English language teaching” (2026), “Enhancing Collaborative Writing Skills Among University English as Second Language (ESL) Learners Through the Utilization of Padlet.” (2025), “Perusall Encourages Critical Engagement with Reading Texts” (2024), “Teaching English as a Second Language in the Midst of a Paradigm Shift: An Exploration of Students’ and Teachers’ Perception of ChatGPT” (2023), and “The Future of Education utilizing an Artificial Intelligence Robot in the Centre for Independent Language Learning: Teacher Perceptions of the Robot as a Service” (2023).



JALTCALL

Trends

Vol. 2 No. 1
2025 JALTCALL
Conference Papers
Special Issue