

# Leveraging GenAI in Creating Audio and Video-Based Flipped Classroom Materials

**Eric FRIGINAL<sup>1</sup> & Jenifer HO<sup>1</sup>**

<sup>1</sup> Department of English and Communication, The Hong Kong Polytechnic University, Hong Kong, China

## Abstract

The move from traditional teacher-centred classrooms to blended and flipped classrooms that are student-centred has called for educators to experiment with new formats of teaching (Friginal & Ho, 2025). Nevertheless, this shift requires educators to develop both pedagogical and technical expertise to design new materials. To better equip educators for this shift, this workshop aims to introduce participants to creating audio and video-based flipped classroom materials using GenAI-powered platforms: ElevenLabs (for audio) and HeyGen (for video). Through interactive lectures, discussions, and hands-on activities, participants will leverage the power of ElevenLabs and HeyGen to create flipped classroom materials that can promote personalised learning and are relevant to students with different needs.

## Outcomes-Based Learning Objectives

By the end of the workshop, participants will be able to:

- Identify and select appropriate GenAI tools for creating engaging, interactive, and personalised audio and video-based materials for flipped classrooms.
- Design (interactive) AI avatars for disciplinary-specific learning.
- Develop best practices for integrating GenAI into designing audio and video-based flipped classroom materials.

## Target Audience

The expected background of participants is instructors teaching in higher education contexts, including but not limited to language-related disciplines. Educators from other disciplines, such as health, social sciences, aviation, and hotel management, will also be able to participate and adapt the pedagogical approaches introduced in this workshop in their own areas of teaching. In general, this workshop will be useful to educators whose disciplines require the simulation of communicative scenarios.

## Topic Relevance

The workshop is closely connected to the sub-themes ‘AI tools and pedagogy’, ‘ethics and integrity’, and ‘incorporating GenAI in learning, teaching and assessment’. The workshop addresses a university-wide initiative to move to a flipped classroom format, as well as aligning with the university’s call for a higher level of integration of GenAI in learning and teaching activities. The workshop also facilitates discussions on the ethical issues of using GenAI tools.

## Workshop Design

### 1. Workshop Format and Activities: Introduction

Multimodal materials, such as audio and video, have become indispensable resources for English Language Teaching (ELT). The trend of using audio and video in ELT coincides with the latest developments in Generative Artificial Intelligence (GenAI) technology. Not only do GenAI chatbots enable students and teachers to interact with tools such as ChatGPT and DeepSeek to generate text-based responses, but the technology has also developed the capability to produce multimodal outputs such as voices and videos.

GenAI technology has also made great strides in enhancing its interactions with users mediated by virtual avatars that sound and look like human teachers. Not only can these virtual instructors present content in video lectures, but with sufficient training, they can also provide answers and engage in conversations in multiple languages.

This workshop focuses on GenAI’s ability to transform the creation of audio and video-based materials in language education, with a focus on interactive avatars and their implications for English for Specific Purposes (ESP). This includes simulating communication scenarios in various disciplines, such as delivering announcements in a flight cabin or managing hotel service encounters.

We start by reviewing literature on the different ways audio and video have been used to support language learning and teaching, such as providing authentic L2 input to improve pronunciation (Scheffler & Baranowska, 2023), speaking and listening skills (Verdugo & Belmonte, 2007; Zhussupova & Shadiev, 2023), followed by a discussion on one specific way of integrating videos in language education – the use of virtual avatars. The workshop concludes by discussing the potential opportunities and challenges, including ethical issues, of using these GenAI tools in language education and beyond.

To establish rapport among participants, the workshop will begin with a quick round of self-introductions and invite participants to share how they have been using GenAI in their own teaching contexts.

## 2. Main Content Delivery

The workshop discusses the potential of using interactive AI avatars as knowledge partners to provide a personalised learning experience for students. The use of virtual instructors that resemble the appearance of human teachers can streamline the material creation process and ensure consistency in the delivery of materials, especially for courses that last over several months.

Virtual instructors provide multilingual support by translating from the content provided by the human teacher, thus catering to individual learners' needs. This is especially helpful in multilingual classrooms with students of different English proficiency levels. Teachers can also develop materials on the same topic with varying levels of difficulty by adjusting the vocabulary used and the pace of speech. They can provide a script for the avatars to deliver using different accents and a range of voices to match the content.

The recent development of interactive avatars allows them to be trained to converse with users on various topics, taking on different roles such as customer support, therapist, doctor's assistant, fitness coach, and more. This potential broadens the scope of their use, demonstrating that they are valuable not only for language teachers but also for educators in other fields with distinct communicative demands.

Besides making use of the available roles, in some platforms, teachers can customise the avatar by training it with the 'knowledge base' necessary for them to perform their role. The prompt is easily customisable, thus allowing teachers to tailor the 'knowledge base' of the avatar with discipline-specific vocabulary and disciplinary norms of interaction to maintain relevance to learners. The avatar can also provide immediate feedback and tailor its language use to be appropriate to the learners. Teachers can easily update the 'knowledge base' of the avatar based on students' needs.

The workshop will first introduce some background studies on audio and video-based material design and GenAI-assisted material design (Friginal & Ho, 2025) using the mode of interactive lecture. Then, participants will engage in hands-on interactive activities to try creating avatars by themselves. Lastly, the workshop will end with a discussion on the potential opportunities and limitations that interactive AI avatars will bring to ELT and ESP. It is our

intention that the workshop will strike a good balance between drawing on research findings and engaging participants in hands-on activities.

### 3. Interactive Activities

#### **Activity 1: Creating Audio and Video-Based Flipped Classroom Materials**

Participants will learn how to create multimodal flipped classroom materials using ElevenLabs (<https://elevenlabs.io/>) and HeyGen (<https://www.heygen.com/>). To increase engagement, the workshop will also demonstrate how to insert human-like AI avatars into video templates or PowerPoint slides that instructors may already have to make the content more engaging for students.

#### **Activity 2: Designing an Interactive AI Avatar for Specific Purposes**

Participants will engage in designing the ‘knowledge base’ required for AI avatars that can perform different roles, such as customer support, therapist, doctor’s assistant, fitness coach, and more. Participants will be shown how to update the ‘knowledge base’ of the avatar based on students’ needs and disciplinary requirements.

Participants will be guided to design the ‘knowledge structure’ of the avatar according to the following perspectives as suggested by HeyGen:

1. Persona of the avatar
2. Knowledge base
3. Language practice activities
4. Feedback and encouragement, among others.

### 4. Discussion and Q&A

- How do we, as educators, ensure that these tools augment rather than replace human interaction?
- How do we guard against the misuse of cloning technology, such as creating deceptive or harmful content?

### 5. Closing Remarks

The rise of voice and video cloning technologies, once reserved for experts, is now within reach for language teachers at a relatively low cost. Virtual instructors open doors to personalised, immersive, and flexible language learning experiences like never before (Friginal & Ho, 2025). But we must also confront the dangers that come with this transformation. These increasingly ‘human-like’ virtual instructors blur the line between what is real and artificial, raising critical questions about authenticity, ethics, and the potential over-

reliance on technology in the classroom. As these tools grow more sophisticated and powerful, the urgency to define their role in education—and to safeguard their use—cannot be overstated.

After the workshop, participants are encouraged to start using the (interactive) AI avatar function to create flipped classroom materials. Once they are comfortable with the technology, they can proceed to more advanced uses offered by the platform, such as exploring AI cloning to create a digital clone of themselves, or a customised interactive AI avatar that is relevant to the learning outcomes of their courses.

## **Acknowledgments**

Our work in this workshop was supported by the Financial Support for Book Writing 2023/24 of The Hong Kong Polytechnic University.

## **Ethical Declaration Statement**

N.A.

## **Generative AI Policy**

The author(s) used ChatGPT-4.1 during the preparation of this work to refine their writing. They also studied the functions offered by HeyGen and ElevenLabs to prepare the materials for this workshop. After utilising the tool/service, the author(s) thoroughly reviewed and edited the content as necessary and assumed full responsibility for the publication's content.

## **References**

- Friginal, E., & Ho, W.Y.J. (2025). Technology and language learning: Effective podcasting and video-based language teaching practices in the age of generative AI. Routledge.
- Scheffler, P., & Baranowska, K. (2023). Learning pronunciation through television series. *Language Learning & Technology*, 27(1), 1–16. <https://doi.org/10.64152/10125/73520>
- Verdugo, D. R., & Belmonte, I. (2007). Using digital stories to improve listening comprehension with Spanish young learners of English. *Language Learning & Technology*, 11(1), 87–101. <https://doi.org/10.64152/10125/44090>
- Zhussupova, R., & Shadiev, R. (2023). Digital storytelling to facilitate academic public speaking skills: Case study in culturally diverse multilingual classroom. *Journal of Computers in Education*, 10(3), 499–526. <https://doi.org/10.1007/s40692-023-00259-x>