Pedagogies: An International Journal

ISSN: 1554-480X (Print) 1554-4818 (Online) Journal homepage: www.tandfonline.com/journals/hped20

Gen AI-assisted multimodal meaning *Design*: exercising a pedagogic metalanguage of transposition

Maria Christoforou

To cite this article: Maria Christoforou (07 Jul 2025): Gen AI-assisted multimodal meaning *Design*: exercising a pedagogic metalanguage of transposition, Pedagogies: An International Journal, DOI: 10.1080/1554480X.2025.2522884

To link to this article: https://doi.org/10.1080/1554480X.2025.2522884

Published online: 07 Jul 2025.

PEDAGOGIES: AN INTERNATIONAL JOURNAL
https://doi.org/10.1080/1554480X.2025.2522884

Routledge
Taylor & Francis Group

Gen AI-assisted multimodal meaning *Design*: exercising a pedagogic metalanguage of transposition

Maria Christoforou

Language Centre, Cyprus University of Technology, Limassol, Cyprus

ABSTRACT

The unprecedented technological advancements in education, like artificial intelligence (AI), Virtual Reality (VR), and the Metaverse, have implications for literacy education and what constitutes to be literate in the digital age, particularly in foreign language teaching. University students engage in multimodal literacy practices in their public lives through technologies that enable innovative textual practices on social networking sites. However, in English for Specific Purposes (ESP), multimodal opportunities are fewer, with a focus on teacher-orchestrated practices and tailored materials using laptops. ESP teaching traditionally emphasizes skills development, limiting learner agency. Despite the multimodal affordances of Gen AI productivity tools, multimodality remains peripheral in ESP, particularly regarding student-created multimodal resources. This case study examines Gen AI-assisted multimodal compositions (videos) in *English for Chemical Engineering I*. Using a multiliteracies framework, it analyzes how students transform course content through videos, moving from passive recipients to active *designers* of knowledge. The study includes an in-depth analysis of three student videos, instructor field notes from class presentations and semi-structured interviews to explore the students' reasoning and engagement with Gen AI tools. The findings highlight the need for a pedagogical shift, positioning ESP students as active *designers* of course content and advocating for more equitable uses of Gen AI tools as meaning-making resources.

ARTICLE HISTORY
Received 9 March 2025
Accepted 13 June 2025

KEYWORDS
Generative AI literacy; multiliteracies; English for specific purposes (ESP); multimodal composition; design; pedagogy shift

Introduction

We live in a world of complexity and constant transformation brought upon the unparallel technological advancements in the educational field, like artificial intelligence (AI) (Godwin-Jones, 2024), Virtual Reality (Christoforou & Efthimiou, 2023) and the Metaverse (Thrasher et al., 2024). These developments impact literacy in higher education and redefine what it means to be literate in the digital age for university students learning a foreign language. The latter experience new literacy practices in their public lives prompted by the multimodal affordances of Gen AI features in technologies like TikTok and Instagram. Lim and Tan-Chia (2022) reported the case of teenage students

CONTACT Maria Christoforou maria.christoforou@cut.ac.cy Language Centre, Cyprus University of Technology, Limassol, Cyprus

participating in the popular culture of short video creation on TikTok and sharing to connect with their peers in affinity spaces. It is a fact that young people these days engage in innovative textual practices on their social networking sites, for example, editing videos and images, commenting, critiquing, and sharing content. Educationally, however, in the intersection of pedagogy for multimodality and foreign language teaching and learning, one can discern fewer multimodal opportunities for undergraduate students, especially in the field of English for Specific Purposes (ESP). Student multimodal literacy practices in ESP is an under-researched area (Christoforou, 2022; Plastina, 2013) with a focus on more teacher-orchestrated practices and tailored-maid materials with laptop resources that students are expected to follow. Traditionally, ESP pedagogy has prioritized linguistic features, primarily written and spoken forms, giving less emphasis on visual or embodied modes (Hafner, 2025). Parkinson (2019) also highlights how the rise of multimodality in student written text challenges traditional ESP approaches and also emphasizes the need to expand ESP pedagogy to embrace an interplay between written and visual modes, promoting more agentive approaches for meaning-making. The value of embodied modes (e.g. gestures) in the teaching and learning of ESP spoken discourse is also discussed by J. C. P. Lee (2024), suggesting that such features have often been overlooked in traditional approaches.

ESP emerged in the 1970s alongside the development of communicative language teaching. Its primary goal was to address the linguistic needs of people preparing for specialized academic or professional fields. ESP sought to diverge from traditional language teaching methodologies, which prioritized the memorization of grammatical structures and vocabulary. Instead, it emphasized the functional application of language as a means of communication (Kern, 2000). Incorporating multimodality in ESP has recently started gaining attention, but there remains a need for further research. In the context of a predominantly visual culture in *English for Fine Arts*, and the integration of digital media, ESP students engaged in multimodal modes of meaning-making and communication (Christoforou, 2022). This shift allowed learners to create their own texts and acquire knowledge through multiple representational forms (Jewitt, 2005). Moreover, the proliferation of digital media has led to the exploration of multimodality in high immersion VR environments at a tertiary education level (Christoforou & Efthimiou, 2023). The multimodal affordances of the VR application helped students move beyond traditional linguistic representations of meanings and offered a more experiential way of interacting with course content through simulation and a sense of presence. Research on multimodality and ESP advocates for a semiotic approach rather than an exclusively linguistic analysis (Paltridge, 2013) since ESP teaching revolves around the development of traditional skills, not leaving room for learner agency (Kalantzis & Cope, 2025a). Instead, it focuses on learning in formal instructional designs (Reinhardt & Thorne, 2011). As a result, literacy education for the digital futures of ESP students must extend beyond the reliance on logocentric texts to pave the way for more diverse modes of meaning-making in their field of study. This implies that instructors and students need a metalanguage (Mills et al., 2023), a language for discussing the meaning-making resources of various modes and for supporting the creation of multimodal artefacts (Lim & Tan-Chia, 2022) in their subject area.

As noted earlier, research on multimodality and ESP is still scarce, particularly in the light of the growing influence of digital media, and requires a more in-depth exploration.

With regard to AI tools in ESP, recent studies (Silitonga et al., 2024; Waly & Zakiyyah, 2024) have begun to explore AI integration in ESP instruction with tools like AI chatbots, ChatGPT, Gemini, etc., highlighting their potential to support language learning and content development. However, these studies do not explicitly focus on multimodal student compositions or the development of student agency through the use of AI-assisted multimodal design. In contrast, students in the present study use GenAI tools to create multimodal artefacts. The study emphasizes the role of such tools not merely as content aids, but as meaning-making resources that can expand students' agentive roles in shaping and communicating disciplinary knowledge. The current case study aims to fill this gap by focusing on multimodal transpositional meaning (Lim et al., 2022) through Gen AI technologies and the concept of *Design* as a theoretical framework to analyse the multimodal compositions of ESP students in a public university in the course *English for Chemical Engineering I*. Multimodal texts are implemented in the afore mentioned ESP course (YouTube videos, website articles, etc.), aiming to enhance students' skills, for example, writing and rote-learning of specialized vocabulary, and to extend linguistic knowledge in the subject of their studies. The use of Gen AI technologies (creation of computer-generated comprehension questions, ChatGPT images related to course content, etc.), however, are solely prepared by the instructor to eradicate "cheating" (Kalantzis & Cope, 2025a) from students. Framed by the need to explore the implementation of *design* in students' multimodal compositions in an attempt to deviate from conventional cognitivist orientations to ESP, this case study is guided by the following two research questions: a) What kind of represented meanings emerge from the students' metalanguage of transposition? and b) In what ways do Gen AI-assisted tools impact students' *redesign* processes in the ESP task?

Literature review

The multiliteracies manifesto, introduced by the New London Group (NLG, 1996), came as a response to the fast-changing social and economic conditions that prevailed in the 1990s. This group of 10 scholars recognized that traditional literacy, which focused on reading and writing in standardized forms, was no longer sufficient in the increasingly diverse societies entering the 21st century. Instead, they proposed a broader view of literacy that would reshape the pedagogical boundaries of the literacy curriculum, leading to more transformative meaning-making opportunities for students. In fact, the drastic changes in communication, globalization, and multimedia technologies created the need to envision new directions in the teaching and learning of literacy (Mills, 2011), embracing multiple modes for meaning-making such as linguistic modes (written and oral), visual representations (i.e. images) (Kress & van Leeuwen, 2006), gestures, audio elements, space, etc. that were derived from the affordances of the new media.

The multiliteracies approach has been instrumental in English as a foreign language (EFL) contexts, enhancing learners' engagement with the lesson and language development. Recent systematic reviews on multimodality in English language learning within higher education contexts confirm the growing transformation towards pedagogies aligned with the multiliteracies framework. Rahmanu and Molnár (2024) conclude that the implementation of multimodal techniques in higher education reflects a broader transformation in educational paradigms, preparing students with the communicative

and meaning-making skills essential for modern professional environments. Following a multimodal problem-based approach in a Taiwanese EFL context, S.-Y. Lee et al. (2019) engaged their students in recursive reading and writing processes to create multimodal texts, highlighting the pedagogical value of a multiliteracies approach in EFL contexts in promoting meaningful language use, critical thinking, and learner agency. The multiliteracies framework was also used for enhancing students' intercultural competence and critical thinking through the use of musical input (Fernández-García, 2024). The study highlighted how integrating music and multimodal resources can enhance students' intercultural competence and strengthen their connection with texts from the perspective of global citizens. Further showing the versatility of the multiliteracies framework, Karimi et al. (2023) explored the affordances of immersive virtual reality (IVR) (embodied cognition, presence, agency, and contextualization) for the development of multiliteracies for situated language use. Key principles of the multiliteracies pedagogy such as collaboration and critical engagement were fostered in the virtual context that extended traditional text-based methods. More recently, in the light of digital transformation in education via GenAI tools, Kalantzis and Cope (2025b) argue about educational justice in fostering student agency, by positioning multiliteracies not just as a set of communication skills but as tools for accessing opportunities and harnessing students' potential for meaningful participation in education. They also argue that, even though the generation of a text by GenAI tools follows the *design* principle in formulating a new, meaningful text, it cannot truly innovate or generate new meanings beyond historical forms. Therefore, a multiliteracies approach is essential for learners for interpreting and *designing* AI-generated texts (Cope & Kalantzis, 2024).

The *design* process

The *what* of multiliteracies focuses on the central role of student agency in shaping their own learning experiences, using various semiotic modes of meaning representation through the process of *design* (see Table 1).

The concept of *Available designs* refers to the recognized and known patterns that individuals use to create meaning within a particular culture or context. In this case study, this responded to the findable patterns of meaning students carried with them to construct meaning when represented with the course material. *Designing* represents the transformative work students engage in, using the *Available designs* to represent the world to oneself or others. This process pertained to the students' transformative process of using Gen AI-assisted tools to represent their own course meanings while working on their compositions. The final step culminates in the act of the *Redesigned*, where both the world and the individual undergo change. The multimodal compositions

Table 1. *Designs* of meaning (adapted from Cope & Kalantzis, 2009).

Available designs	Accessible and retrievable resources one has for constructing meaning, shaped by specific patterns and conventions of meaning-making
Designing	The process of creating meaning working on/with Available Designs, using them to represent ideas about the world to express others' interpretations of it, whether for oneself or an audience.
The redesigned	The outcome of the transformation is the world transformed, either as new Available Designs created to represent the world differently, or the meaning designer who, through the act of Designing, has undergone transformation (personal growth and learning).

acted as the new texts of the *Redesigned*, which exerted a dynamic play of student subjectivity (Cope & Kalantzis, 2009). Finally, a design-theoretic approach views learning as a form of communication, emphasizing both the maker's and receiver's agency (Kress & Selander, 2012). This means that the initial sign serves as a prompt for the receiver who interprets it according to their own interest and using their own semiotic resources. This is how new signs of representation of understanding the world are created.

The grammar of transposition

Bill Cope and Mary Kalantzis (members of the New London Group) have significantly expanded the theory of multiliteracies within the past 30 years (2009, 2022) and the dynamic relationship between meaning, *design*, and social context. Their recent *Grammar of transposition* of multiform meaning introduced a conceptual shift, replacing the word *mode* with *form* and *multimodality* with *transposition* in the era of ubiquitous digital media where text, image and sound exist in the same form. They argue that meaning is not merely conveyed through static forms but is continuously reshaped across different forms and contexts.

With their framework of transpositional grammar, Kalantzis and Cope (2022) support that the emphasis on multimodal or multiform meaning does not diminish the central role of language in the process of meaning-making. Rather, it acknowledges that language is embedded within an ensemble of other semiotic forms (i.e. text, image, body, speech, etc.) and the orchestration of meanings often involves movements across these forms and/or an interplay of these forms (Lim et al., 2022). Kalantzis and Cope (2022) introduced the term *function* to examine the shifts in the patterns of meaning and the different purposes that every meaning serves. The model builds upon Michael Halliday's (1978) functional linguistics and notion of language as social semiotics, which originally identified three metafunctions: ideational, interpersonal and textual, only to be expanded by Cope and Kalantzis (2020) with the addition of two more metafunctions, *context* and *interest*. They support the idea that in this era of technological proliferation, pervasive digital media and their materiality can constrain or enhance the nature of semiotic forms in the communication of meanings; therefore, meanings should not be seen in isolation but in their imminent transposability. Regardless of its form or combination of forms, the five functions are always present in meaning-making (Lim et al., 2022). *Reference* addresses Halliday's ideational question, "What is this about?" – this text, this image, this sound, this speech or any combination of these. The second function, *agency*, address the interpersonal question of who or what is performing this action and their role as meaning-makers in the meaning process. *Structure* addresses Halliday's textual metafunction and answers the question, "How do the meanings cohere?", referring to the coherence of the discourse and how it is organized (Lim et al., 2022). *Context* responds to the question, "What is this connected with?" which aims to determine how meanings fit into the larger world of meaning. Finally, *interest* elucidates the underlying motivation for meaning-making, whether human-driven or naturally occurring, and answers the question, "What purpose does it serve?".

The transposable and diverse nature of meanings is what makes them fluid, a fact which stems from globalization and digital technologies, like Gen AI. Even though the materiality of forms may influence the type of meanings that emerge, and even though Gen AI works predominantly with written forms (Cope & Kalantzis, 2024), every moment of meaning enacts all five functions.

Materials and methods

Context

The study took place during the fall semester 2024 in the ESP course *English for Chemical Engineering I*, a three-hour-per-week course during the students' first semester of studies in a tertiary education institution in Cyprus. The course is specifically designed to meet the linguistic needs of university students in the field of Chemical Engineering and to develop their English proficiency to an academic level equivalent to B2 on the Common European Framework of Reference (CEFR) for languages. Students of this ESP course could access the course material through the Moodle platform, which is the institution's Learning Management System (LMS). Some of the basic thematic units of the course were: *Introduction to basic concepts of Chemical Engineering*, *Introduction to batch and continuous manufacturing processes*, *A day in the life of a Chemical Engineer* (basic responsibilities), *Intermediate products vs finished products*, *Virtual Reality in Chemical Engineering*, etc. As the institution's official LMS, Moodle predominantly provided students with written texts (Figure 1)

for studying (online articles, excerpts from books), but they also studied content from YouTube videos (Figure 2), static images, and other multimedia resources.

English for Chemical Engineering I was conducted in a medium-sized computer lab which allowed all students to have a laptop at their disposal for the purposes of the

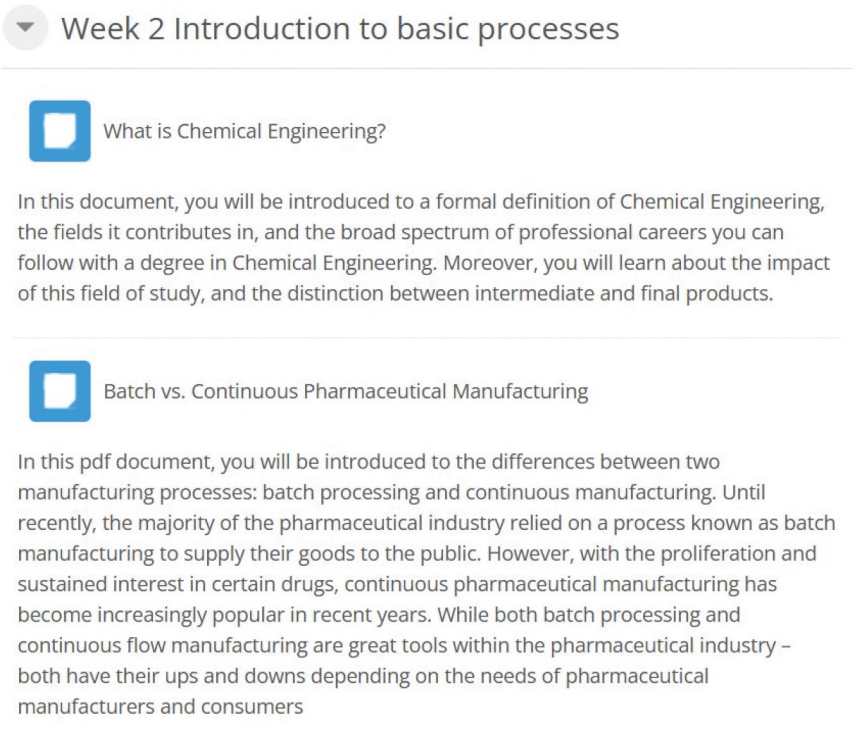

Figure 1. Written texts as basic course material on Moodle.

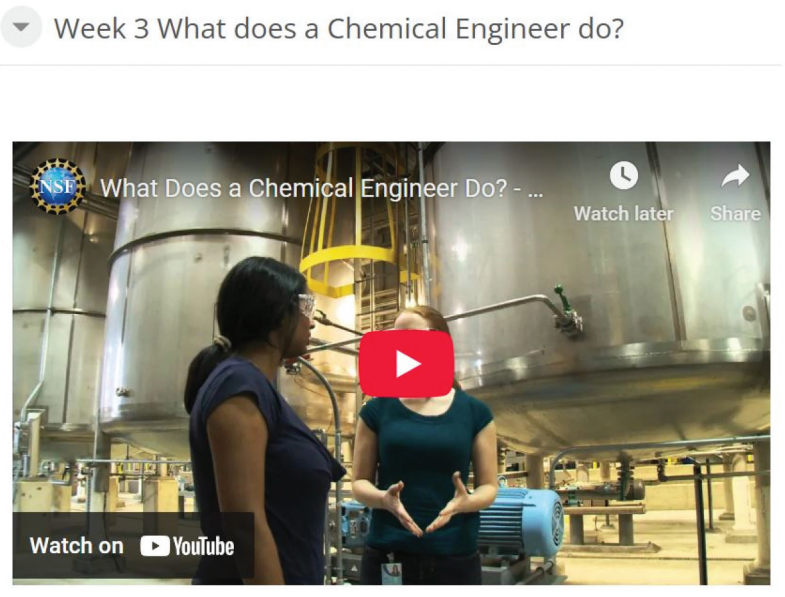

Figure 2. Implementing YouTube videos as a learning resource.

lesson. Concerning the spatial arrangement of the lab, the layout came with a linear, row-based arrangement with computers being placed in straight rows facing the direction of the instructor. Due to the limited use of space, the instructor could not easily walk towards the students close to the wall and students could only interact with the student to their left or right.

The multimodal composition (video)

The students' semester project in the current study asked students to create their own video (Figure 3) as a way of reproducing the thematic unit of their interest.

Lim and Tan-Chia (2022) refer to digital multimodal compositions as learning processes that enable students to transfer their learning to new contexts, involving embodied meaning-making opportunities. Moreover, the instructor selected the video as a medium that would enhance students' multimodal meaning construction (Lim, 2020). In this study, students worked in groups of two or three, and they were free to choose between the topics, the types of software, and the semiotic modes, using English as the language of delivery.

Participants

Six participants, aged 18–20 years old, all female, were selected out of a larger group of 18 students. Given that only three of the students in the class were male, the predominance of female participants in the data reflected the class composition. The selection of an all-female participant group was, therefore, not intentional, nor was it influenced by gender considerations. Assumed names have been used throughout

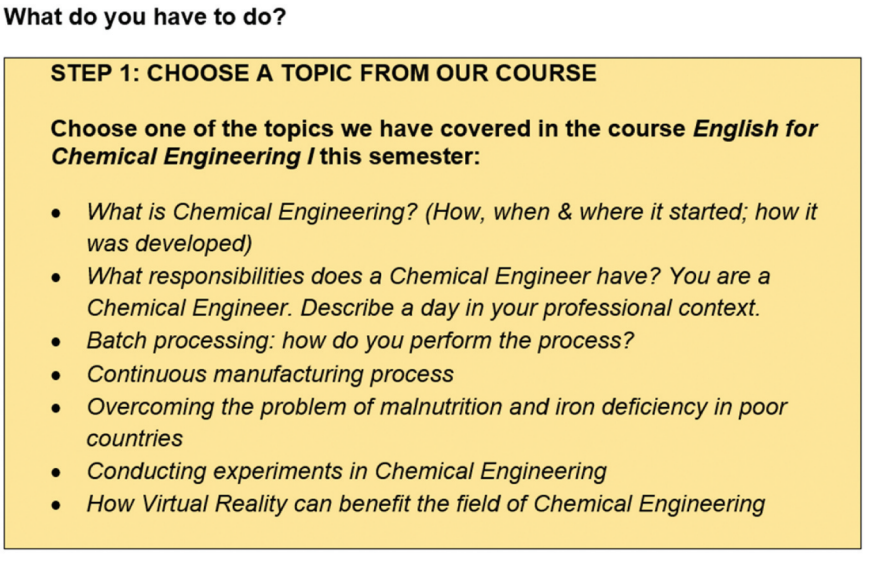

Figure 3. A list of all the topics of the semester project.

the paper to maintain the anonymity of the students. The six participants were purposefully selected (Creswell, 2015) based on specific criteria, including their high level of creativity in multimodal compositions, particularly in how they approached the topics, their technical skills in areas like filming and editing (Lim & Tan-Chia, 2022), and their effective use of transpositional grammar strategies to represent course content across different forms. These criteria made them well-suited for an in-depth exploration of design processes facilitated by GenAI tools. In the context of this case study, the aim was to explore the students' *designing* through a metalanguage of transposition across different forms and functions in Gen AI-assisted multimodal videos and it was decided that the compositions of these six students were more pertinent for the purposes of this research. Moreover, to mitigate potential bias stemming from the instructor also being also the researcher, students were informed that participation in the study was voluntary and would have no impact on their grades in the course, ensuring that that they did not feel they were working in an evaluative environment.

All participants were of Greek-Cypriot origin and had studied English as a foreign language at school. They used TikTok and Instagram for entertainment and were active in social media sites. At the time of the study, various GenAI tools, for example, ChatGPT, Grammarly, Gemini, and Animaker, were widely available and accessible in their free versions to be used in educational contexts. For the purposes of this study, no specific GenAI tools were recommended or promoted by the instructor. Students were free to choose which tools to use, based on their existing familiarity and preferences, but they were encouraged to rely on free versions rather than paid subscriptions.

Table 2 shows the three groups of students, their chosen topic of the multimodal compositions and the Gen-AI tool they have used for their work.

Table 2. Topic and Gen-AI tool for the students' multimodal compositions.

Groups of participants	Topic of the multimodal composition	Gen-AI technology used
Group 1: Conny & Micha	*What is Chemical Engineering?*	ChatGPT, Animaker
Group 2: Helena & Zia	*How the term Chemical Engineering was coined*	ChatGPT, Filmora
Group 3: Georgina & Mariana	*Creating a lipstick through a batch process*	Animaker, Adobe Animate, ChatGPT

The collected data

The present study employs a qualitative method of data collection: multimodal compositions (videos), the instructor's field notes from the students' class presentations and semi-structured interviews conducted in-person with the group of students at the project's conclusion. Qualitative methods enable the development of more exploratory and in-depth models of research (Robson & McCartan, 2016). Additionally, small-scale research like the present one aims to represent itself and understand the study phenomenon. It does not require generalizability of results and representativeness of a wider population.

The students' multimodal compositions were sent to the instructor via WeTransfer, a computer file transfer service that allows users to share large files without registering for an account. The instructor's field notes of the students' class presentations were organized into four thematic sections: content, creativity, challenges met, and presentation/audience engagement. The semi-structured interviews were conducted in the participants' native language, Standard Modern Greek, to elicit a richer and more diverse range of responses. The interviews were audio-recorded, transcribed verbatim in Greek, and manually translated into English by the researcher.

Data analysis

The videos were analysed deductively (Saldaña, 2021, p. 40), attending to the semiotic forms that the students used in their compositions, with the codes driven by the theory of *transpositional grammar* and the five meaning metafunctions (Kalantzis & Cope, 2022): *Reference*, *Agency*, *Structure*, *Context*, and *Interest*.

The field notes and semi-structured interviews were manually coded by the researcher using thematic analysis (Saldaña, 2021). Credibility was checked through "investigator triangulation", involving a process of an external researcher auditing the data collection and analysis (Creswell, 2015). Table 3 shows the data collection sources and methods of analysis.

The questions used in the semi-structured interviews were meant to encourage students to provide feedback on their Gen AI-assisted video, providing more inclusive input (Tzirides et al., 2025) on the reasoning behind their choice of AI tool and their pedagogical engagement with it. The questions that were asked are the following:

Table 3. Data collection and analysis methods.

Data-collection sources	Data collected	Methods of analysis
Students' multimodal compositions	The content of the three videos	Multimodal analysis
Instructor's field notes	The content of the notes	Thematic analysis
Semi-structured interviews	The transcribed content of three interviews	

(1) What technologies do you use in your daily life to create any content?
(2) Have you ever used technologies for educational purposes before your English course?
(3) Which topic have you chosen for your multimodal composition?
(4) What AI tools did you use for your work?
(5) Why did you choose these specific tools?
(6) In what way or ways do you think you showed your topic through these specific tools?
(7) How did you ensure the authenticity of your work so that it would not be considered as cheating?
(8) Regarding the topic you chose, to what extent do you think it changed the way you and others who see it understand it?

Results

This section presents the findings of the analysis, which are divided into two subsections: multimodal analysis of the students' videos, and thematic analysis of the instructor's field notes and the semi-structured interviews.

Multimodal analysis of compositions

This section analyzes students' Gen AI-assisted multimodal compositions (videos). The excerpts presented here are drawn directly from the video content itself, not from interviews or reflections.

Video 1: Conny and Micha - What is Chemical Engineering?

The first group consisted of two female undergraduate students, Conny and Micha. To develop their video presentation on the topic *What is Chemical Engineering*, they utilized the Gen AI animation tool, Animaker. Through this technology, they represented themselves using avatars wearing professional attire, and maintaining a formal-looking digital presence throughout the video.

Reference. The students conveyed their designed representation of what chemical engineering is, and the industries it applies to (i.e. energy production, healthcare, the environment and food production). It is interesting to see how the students used a combination of forms to realize *Reference* from the very beginning of the video. The presentation started off with an introduction in which they identified themselves as first-year students and indicated the name of their public university. Figure 4 shows the students' interplay of forms to introduce viewers to the topic.

Agency. To realize *Agency*, the students actively participated in the meaning process through AI-generated avatars embedded within Animaker, taking on professional roles that were visually constructed through the avatars' attire (lab robes). In the video, the students dressed their avatars in formal clothing, visually signalling professionalism. The lab robes also worked as signifiers, representing work in, both, laboratory and industrial settings.

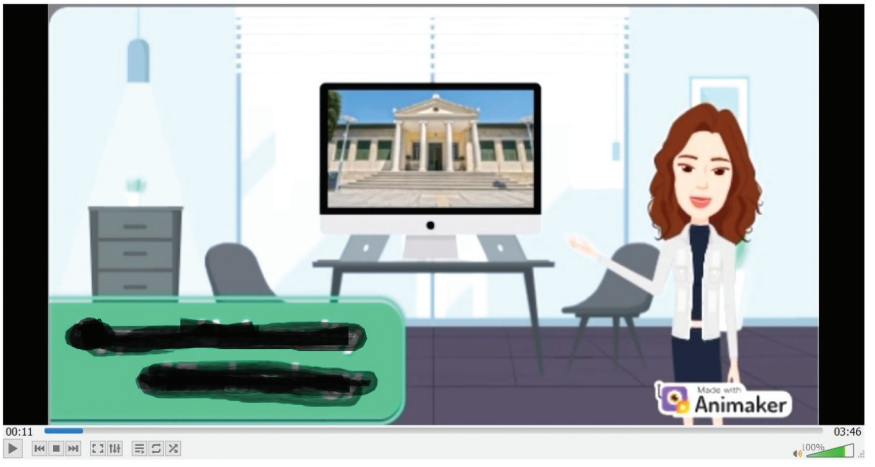

Figure 4. The semiotic interplay of forms in the introduction of the first video.

Structure. The scenes of the video worked coherently through transpositional orchestration, where different forms interacted to establish cohesion. Meanings hung together via the students' real voice which was overlaid onto the video (speech) and which aligned with the visual elements (AI-provided or generated images) to strengthen meaning, introductory music (sound), avatar movements (body), spatial proximity to embedded objects in the video (space), and students' written prompts (text).

Context. The students situated the Gen-AI assisted meanings related to Chemical Engineering to a broader context than the educational one (i.e. university). The conceptualization of their conveyed meanings extended beyond the immediate topic, encompassing AI-generated video production that tied to the students' existing literacy practices in digital content creation. It is a fact that the creation of the AI video did not only connect with the students' disciplinary knowledge. It also intersected with their everyday digital communication practices and multimodal familiarity in terms of content creation.

Interest. In terms of purpose, the students' AI-generated video aimed to inform viewers about Chemical Engineering and what it entailed. While the video aligned with the students' academic objectives (motivation to do well in their English course), it also sought to stimulate the viewers' curiosity and interest to prospectively explore further the field through the strategic use of simplified visual forms (realistic representations of their university), lab coats (objects), textual elements (written points in bullet points), authentic settings (i.e. space in a factory or in a conference room), etc.

Conny's use of metalanguage in the video to *redesign* the topic for viewers is evident in her choices, such as integrating music, selecting a relatable avatar, and incorporating images, to keep viewers engaged. Even though Gen AI generally prioritizes written text as a technology, it can work with other forms of meaning when provided with other textual forms (Cope & Kalantzis, 2023, p. 10). This allowed her to assign meaning through multiple

forms in a specific context (in this case, a professional conference room). In the video, the student animated her avatar to use body gestures and point at the slide on the projector, synchronizing this visual action with her real voice (speech) overlaid onto the video: "We are first-year undergraduate students in the field of Chemical Engineering at the (name of university). Today we would like to talk to you about Chemical Engineering". Speech is aligned with the avatar's movements (body) to enhance authenticity and complement the visuals. In terms of space, the student's avatar is positioned in a realistic conference room setting, utilizing spatial elements such as proximity to a table and chairs (objects), mirroring a real-life presentation. This strategic use of space enhances the authenticity of the virtual presentation, reinforcing naturalistic communication dynamics. In terms of text and image, Conny integrated a realistic photograph of her university, featuring its actual name in written form (wiped out here), digitally recontextualizing the learning environment to bridge the physical and virtual classroom space.

For the second part of the video, Conny and Micha informed viewers about the important role chemical engineers have in converting raw materials into products and the various industries they can work in (*Reference*).

In Figure 5, we can see Micha orchestrating her meanings and taking on a more agentive role to realize *Structure* (establishing coherence in the video) and *Context* (placing meaning within the social context viewers will understand). The reason is that the AI images of Animaker (*a wind turbine, a microscope*, *a factory*, and *an apple* and the embedded objects in the background wallpaper do not resonate accurately with the intended meaning of *industries* and do not scientifically portray the topic. Rather than relying solely on the visuals provided by the AI software, she engages in a transpositional shift across meaning modes and meaning functions (Lim et al., 2022, p. 6) to enrich the scientific representation of the topic, predominantly through her own verbal explanations (speech derived from written text). She associates specific images with corresponding industries, demonstrating her interpretative engagement with the visuals. When a wind turbine is depicted, she identifies it with the energy industry and also links the image of

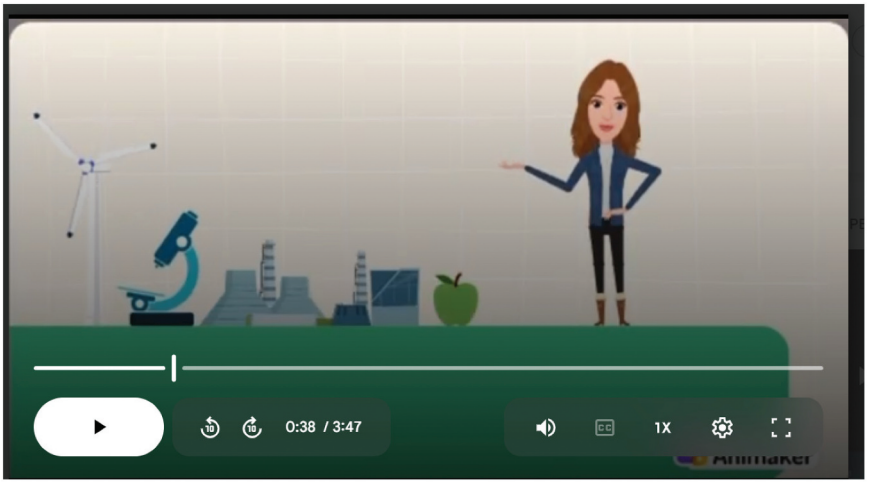

Figure 5. Micha realizing *Structure* and *Context* in her multimodal composition.

the microscope with healthcare. As for the last two industries, she associates the image of the factory with environmental protection and interprets the apple as a representation of the food production industry. This process highlights the fluidity of meaning in AI-provided or generated images and the role of verbal reinforcement in clarifying intended concepts. This interplay between image and language allows her to participate in the meaning process more effectively (*Agency*) by bridging the gap between the AI-provided content and her personal understanding and interpretation. By doing so, she transforms the digital representations into a more nuanced and contextualized learning experience, ensuring that her intended message is clear and aligned with her broader world experience (Kalantzis & Cope, 2025a).

Video 2: Helena and Zia - How the term Chemical Engineering was coined

The second group consisted of two female undergraduate students, Helena and Zia. To develop their video presentation on the topic *How the term Chemical Engineering was coined*, they used the Gen AI video editing software, Filmora. Through this tool, the students edited and connected scenes, they added music and special effects, and adjusted transitions to enhance the flow of their narrative. Diverging from the previous avatar-based video, Helena and Zia chose a more unique representational approach to their topic by actively embodying historical roles and posing as the founder of Chemical Engineering, George E. Davis and his wife Laura Frances Miller (*Agency*) in a humorous way (Figure 6).

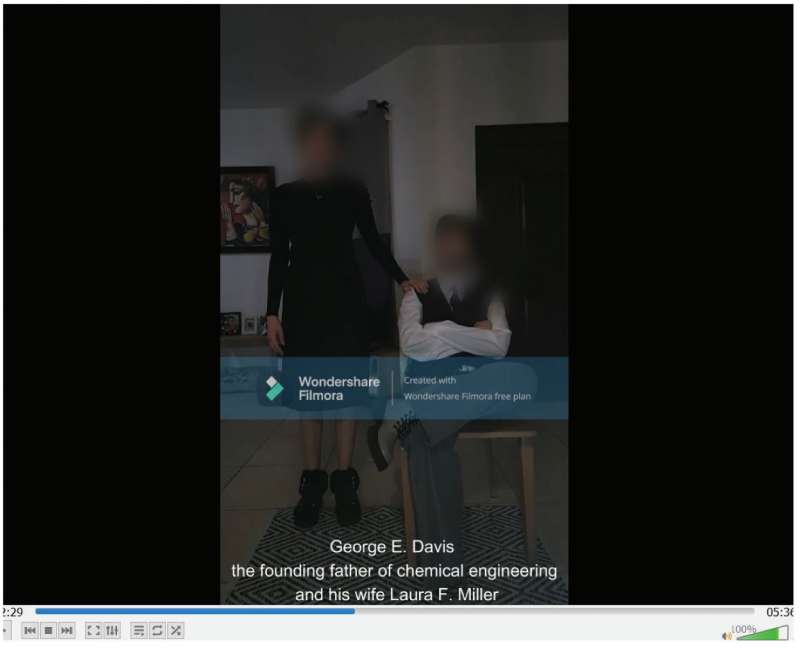

Figure 6. Students posing as the founder of Chemical Engineering and his wife.

Reference. The students used a pedagogic metalanguage as a resource to represent and communicate their own version of how the term Chemical Engineering was coined, adapted from the printed text they had studied in class (Lim & Tan-Chia, 2022). It is important to mention that, through the video, the students expanded the potential for meaning-making since the circumstances under which the term was coined were not taught to the students.

Agency. As mentioned, Helena and Zia engaged in a dialogic re-enactment of a family scene between the founder of Chemical Engineering, George E. Davis, and his wife that led to Davis coining the term "Chemical Engineering". Their approach foregrounded agency since they remade and transformed more creative meanings (Cope & Kalantzis, 2000) through their own lens by wearing historically accurate attire (objects), making use of space (their own kitchen) to act out the roles, and writing their own dialogues for Davis and his wife (text).

Structure. The transposition of meanings across the material forms to realize *Structure* resulted in a different representation of the topic, contrasting with the more formulaic approaches to text in the ESP course, which typically follow predefined templates for text organization. In the students' multimodal composition, meaning was established through multiform movement. For written text, the students used their own constructed dialogues, while, at the same time, vocal delivery (sound and speech) was strategically employed to emphasize key points and convey nuances such as the wife's disregard or dissent when it came to her husband's long hours of work. Their semiotic activity was further reinforced through their historically accurate attire (objects) to ensure a more realistic representation of the topic in terms of time (1881), place (their place of residence), and theme (discussing a batch process that led to Davis' coining of the term). Figure 6 also visually depicts their body posture which remained formal and rigid (body movements) to align with the rest of the forms. Finally, all previous forms were reinforced as the students synchronized various forms in their video by adding music at the end of the video and introducing images containing written text to emphasize key information about Chemical Engineering, effectively concluding the narrative of Davis' discovery.

Context. The students extended the Gen AI-assisted meanings related to the coining of the term Chemical Engineering beyond the academic focus. They enriched their video production with creative and technical competencies such as editing the software, positioning and handling of the camera (Lim & Tan-Chia, 2022, p. 74), all of which allowed them to establish a relationship with historical figures from their field of study. Their choice of clothes (objects), make-up, and posture (body) were the predominant semiotic forms that held a central position in meaning making. This is because these forms activated sensory input (sight, sound, sense of spatial awareness) from the audience that speech and text could not achieve alone (Kalantzis & Cope, 2022, p. 38) and because they aligned accurately with the setting (space) and historical interpretation of the scenes.

Interest. According to Lim and Tan-Chia (2022), *interest* relates to the emotions of students, the social impulses, and the reasoning behind their creations. In Helena and Zia's case, understanding the purpose of the multimodal composition is marked by how the purpose is shaped across multiple forms of meaning. Their dialogue (written text) conveyed a layer of subjectivity with a humorous predisposition that deviated from objective facts taught in the lesson, particularly at the beginning of the video where Davis' wife questions him, saying "You were with those so-called chemists?". This evident scepticism towards the chemists' work also revealed the wife's perception of them as foolish for dedicating so much time on chemistry. Moreover, *interest* is primarily transposed from the AI-embedded subtitles (written text) across Helena's embodied action (body movements), as she portrays Davis rising from his chair and having an epiphany upon discovery the term "Chemical Engineering". This conveys the students' potential belief in how great minds in their field of study conceptualized scientific terms and potentially achieved career breakthroughs.

Video 3: Georgina and Mariana - Creating a lipstick through a batch process

Georgina and Mariana were the members of the third group and together they created a video on the process of creating a lipstick through batch processing, using the Gen AI tool, Animaker, and Adobe Animate. The students generated a logo for their cosmetics company (Figure 7) using Adobe Animate, enhancing the authenticity of their production process.

Reference. *Reference* was realized through a series of embodied actions (body), with the students using a metalanguage to describe their batch process for lipstick production. Through their embodied actions (i.e. showing the ingredients to the audience and

Figure 7. The students' logo of cosmetics company made via Adobe Animate.

labelling them, testing the final product themselves, asking viewers to subscribe to their channel, etc.) and the portrayal of realistic ingredients (objects), Georgina and Mariana represented batch processing in an authentic *How-to* YouTube video format.

Agency. The students assumed a highly agentive role as experts in the field (Figure 8), reinforcing their image through the use of lab coats, real ingredients for the process (objects), and the Gen AI-generated logo of their business, all of which enhanced their credibility in the eyes of the audience.

Moreover, they presented each ingredient (i.e. beeswax, beetroot, shea butter, etc.) with a written representation on the screen (text), ensuring a comprehensive, and professionally structured presentation of the batch process.

Structure. Coherence was achieved through an orchestration of multiform meanings functioning simultaneously to structure a beginning, a middle, and an ending to the batch process (Kalantzis & Cope, 2022). Their transformative work did not only focus on the sole recreation of the lipstick manufacturing, but on reconstructing a meaningful portrayal of the batch process within the context of the digital culture of YouTube, engaging it with viewers and subscribers. At the outset of the video, viewers could see the AI-constructed logo of the girls' cosmetics company, accompanied by music embedded in Animaker (sound). By incorporating the digitally designed logo into their video, the students established a recognizable brand presence, which, together with the title "Make a lipstick with Georgina and Mariana," reinforced their authoritative role and revealed their expert identities to the viewers. Connection was also established verbally via the students' opening phrase in the video, "Hey guys, welcome back to our channel. Today we are going to make a lipstick, using ingredients that you can easily find". This phrase is aligned with the conventional discourse that is internally associated with *How-To* YouTube videos that one can watch online.

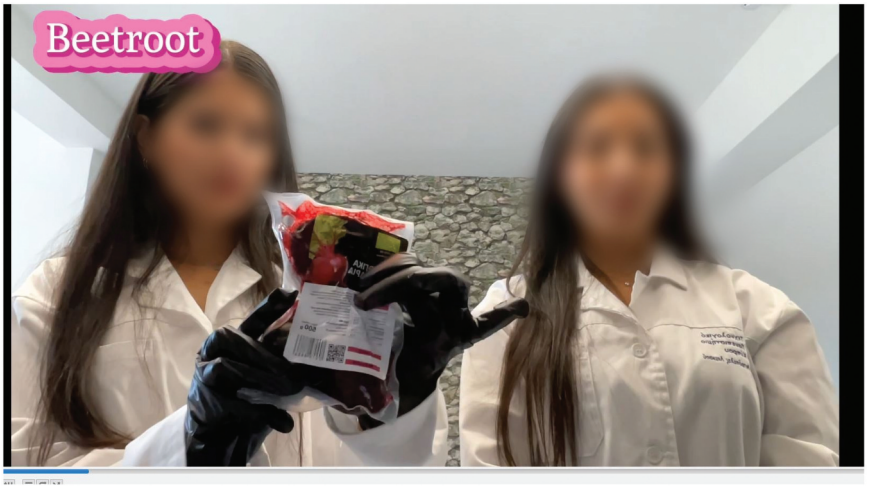

Figure 8. Georgina and Mariana as experts in the field in a *How-to* YouTube video.

During the middle of the video, the students initiated the lipstick process by orally presenting the ingredients viewers would need to use. For instance, when introducing beeswax, Georgina stated, "This is beeswax. It acts as a solidifying and thickening agent and provides the lipstick with a firm structure". Such verbal prompts added credibility to the students' meanings, as they were positioned as engineers with knowledge and expertise. This enhanced the reliability of the process in the eyes of the viewers. Beyond the oral descriptions of ingredients and procedural steps, the students used Filmora to layer their metalanguage across other forms. This included the written names of the ingredients, allowing viewers to visually associate the words with the images, as well as hand movements (gestures) for showing plates, knives, etc. (objects). Additionally, the students edited their footage to add meaning through visual, auditory, and textual forms. For example, for the actual process, they synchronized their hand movements (Figure 9) with background music and included on-screen written prompts, enhancing the multimodal composition.

At the conclusion of the video, prior to revealing the final product (lipstick), the students switched to an AI-generated avatar image of themselves (Figure 10), depicting two elegantly dressed female figures. This visual representation symbolically aligned with themes of beauty and sophistication, resonating with the intended audience who would attempt to create this lipstick. The transition to the students' actual representation continued to create meaningful processes as they no longer wore their lab robes. Instead, they changed to casual attire, holding the lipstick and smiling for the success of the end product (body movements and object), displaying also their website (text) at the bottom of the screen (www.gmcosmetics.com). This strategic integration of the website established meaningful associations between the student engineers and the audience, the scientific process, and the perceived expertise of the students, reinforcing their professionalism.

Figure 9. Enhancing the batch process with hand movements.

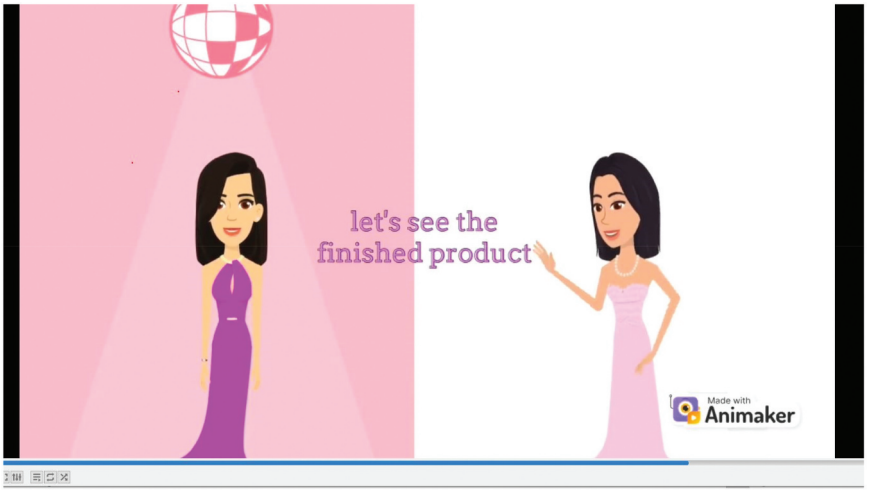

Figure 10. Switching to two female Gen AI avatars.

Context. The interplay of forms and shifts in meaning collectively contributed to the meaningful production of a *How-To* video, similar to the videos played on YouTube with viewers and potential subscribers. Using Animaker and Adobe Animate, the students crafted their videos in a familiar and interactive communicative style that mirrors the contemporary digital culture of platforms like YouTube and TikTok to target their audiences.

Interest. *Interest* was realized through the students' motivation for creating the video, driven by an intent to exemplify a batch process not usually found in books. This implied an objective to not only illustrate the technical aspects of the batch process but also to present the making of a lipstick (object), a product that resonates predominantly with female audience, including the students themselves. Their multimodal composition reflected a blend of objectivity and subjectivity. From an objective standpoint, Georgina and Mariana conveyed the functional properties of the ingredients through verbal explanations while simultaneously demonstrating how they were used in the production process (using body movements and gestures like cutting and stirring). This reinforced an understanding of how each component contributed to the overall experience. However, their composition also incorporated some subjective elements as their choice to describe the creation of a lipstick aligned with their personal interests and aesthetic preferences. Furthermore, their decision to integrate an AI-generated logo (visual representation) for the company established more credibility and presence as scientists, along with their website (textual element), all of which enhanced their professional presence with the audience.

Thematic analysis of the instructor's field notes and semi-structured interviews

Thematic analysis of the instructor's field notes and semi-structured interviews offered insights into how the Gen AI-assisted tools impacted the students' *redesigned* processes in

Codes and related excerpts

Efficiency and structural support	Engagement through Entertainment	Skepticism in AI Credibility	Increasing student agency
I wrote in English "Make me a structure about what is Chemical Engineering as a presentation" and it gave me some bullet points and told me "At the beginning, give the definition of What is Chemical Engineering", then it divided them into field 1, field 2, field 3, etc., so it helped us a little with the structure in how to talk about each thing because at the beginning it is a little chaotic. There is a lot of information and we wanted it to divide it up a little so that we could develop it. **(Group 1)**	*I said let's do something more fun, to have a better time doing it and to want to do it so that it's not boring.* **(Group 2)**	*And then we wrote our own words and changed some things that ChatGPT translates or connects that don't make sense, so you have to check it later and change some things too (.) ChatGPT brought out some terms that we didn't understand.* **(Group 2)**	*Students feel seen and heard. It is important to see the lesson from the students' voice because it becomes more interactive and it encourages creativity.* **(Group 1)**
With ChatGPT, it didn't make it so chaotic. I think that if we didn't have ChatGPT, we would have thrown our ideas around a bit, mixed up the topics a bit, jumped from one topic to another, whereas now it had a better structure, so I think the rest of the class understood it better. **(Group 1)**	*We wanted to make it more realistic and more fun; to capture our fellow students' attention from the beginning of the video.* **(Group 3)**	*We basically paraphrased. ChatGPT wasn't very familiar with batch processing and what we wanted to do. It wasn't very accurate in what it was telling us, I mean it was more general.* **(Group 3)**	*When students are invited to express their understanding and interpretations, they feel empowered and take ownership of their learning process, which boosts motivation and confidence.* **(Group 2)**
We also had some handouts from class and we had ChatGPT tell us what each thing meant, so that we would know fully what it meant but they would also understand it because it made it simpler. **(Group 2)**			*We completely got into the role of the character and in order to be able to portray them you definitely have to know the story because you can't show something or imitate someone and not know exactly their story so we both looked for similar things and got completely into the mood, to learn what that era was like, and how both characters were connected because it was both androgynous so we had to connect it somehow like that, and I personally remember everything I read and everything I wrote.* **(Group 2)**

Figure 11. Excerpt of researcher's coding document.

the multimodal compositions for *English for Chemical Engineering I*. The process of moving from codes to themes began with the collection of all the excerpts coded under each code. Figure 11 shows an excerpt of the researcher's coding document.

Adopting an inductive analysis approach, and allowing the data to speak for itself (Saldaña, 2021), the study identified the themes as presented in Table 4.

Theme 1: AI as a tool for learning support and engagement

Code 1a: Efficiency and structural support from the Gen AI tools

Table 4. The most common themes that emerged from the transcribed data.

Codes	Themes
• Efficiency and structural support from the Gen AI tools • Engagement through entertainment	AI as a tool for learning support and engagement
• Skepticism in AI credibility • Increasing student agency	Critical thinking and student empowerment in AI use

Code 1b: Engagement through entertainment

Thematic analysis revealed students' perception of AI as a learning support tool that enhanced their multimodal compositions both structurally and creatively. Aligned with the pedagogical possibilities for multimodal literacies that Gen AI introduces (Kalantzis & Cope, 2025a), students used Gen AI tools to *redesign* course content in fun and engaging ways. For example, Group 2 used Filmora "to put scenes together, add images, overlay text, insert music, and add effects between scene transitions", demonstrating their ability to apply a metalanguage to discuss the meaning-making resources of the forms they employed (Mills et al., 2023). The process was described as entertaining and creative (Lim & Tan-Chia, 2022, p. 71), reflecting the students' comfort and familiarity with AI tools in their daily lives.

Similarly, Group 3 explained, "It made the video more realistic and fun to watch so that our classmates would be engaged from the start." Group 1 emphasized how the tools supported them structurally:

> Without Animaker, we wouldn't have been able to create the animation. It helps people understand better because they can see, hear, and read at the same time - engaging three senses. Without it, we wouldn't have had an animation. And without ChatGPT, we wouldn't have been as well-guided in structuring our presentation to have a clear beginning, middle, and end.

This supports the idea that Gen AI tools serve as scaffolding mechanisms, enhancing both organization and multimodal comprehension (Kalantzis & Cope, 2022). By emphasizing that they could "see, hear, and read at the same time," the students showed how they used the multimodal features of Gen AI tools to enhance their understanding through multisensory engagement.

Theme 2: Critical thinking and student empowerment in AI use

Code 2a: Skepticism in AI credibility

Code 2b: Increasing student agency

While students appreciated the tools' support, many voiced scepticism about the reliability of AI-generated content. Georgina and Mariana noted, "ChatGPT wasn't very familiar with batch processing and what we were trying to do. The information it gave us was too general". Zia added, "Sometimes, ChatGPT's translations or connections didn't make sense, so we had to review and modify them … We were not familiar with some terms it suggested." These comments show that students were not passive users but critically assessed and adapted the content to better suit their needs.

Rather than using AI tools uncritically, students demonstrated a high level of agency, questioning and revising what was offered, effectively challenging concerns around academic dishonesty and "cheating" (Kalantzis & Cope, 2025a). Their active stance

illustrates a process of critical engagement and reflective decision-making in the context of their disciplinary learning.

Discussion

This case study has examined ESP students' Gen AI-assisted transpositional meaning *designs* through the creation of artefacts (videos), using a metalanguage to describe the processes of meaning-making that extended beyond the linguistic form (Lim et al., 2022). The students' metalanguage led to more pragmatic language uses in *English for Chemical Engineering I*, since they communicated their meanings in more contextually appropriate and realistic ways, for instance, Group 2's reconstruction of how Chemical Engineering was coined through Gen AI-assisted scene editing, using a metalanguage to contextualize a historical event which had not appeared in their LMS. Similarly, in the simulation of a batch process, the students of Group 3 had to navigate specific procedural terminology through the transpositional affordances of Animaker. Their use of metalanguage enhanced communicative precision about their lipstick, fostering more applied and pragmatic uses of language. Moreover, the multimodal analysis of the students' artefacts underscores the role of Gen AI tools as semiotic technologies that function not only as meaning-making tools but also as semiotic practices (Lim & Tan-Chia, 2022), evident in the way students' transpositional work reflected practices embedded in their daily digital interactions and cultural experiences (Kern, 2000), such as creating video content, editing, and providing commentary in YouTube-style *How-To* videos. While such practices with Gen AI tools benefit from digital affordances, they can also be adapted to less tech-equipped contexts, as many of these tools provide free versions which students can use. This flexibility can support student agency and multimodal engagement across a range of disciplines, including the sciences, arts, and humanities. Moreover, this study reveals a recontextualization of the learning environment, as the multimodal compositions supported the extension of the classroom space beyond its physical boundaries and broadened the scope of the learning process.

In the context of ESP, the study contributes on two interconnected levels. First, it offers empirical documentation of how ESP students in a higher education context engage with Gen AI tools to support meaning-making and content creation. Their critical engagement, strategic use of multimodal affordances, and revision of AI-generated results provide insight into the evolving literacy practices in digitally mediated contexts. It also sheds further support to the need to reconceptualize meaning-making practices and redesign academic literacy in digital contexts (Christoforou, 2022) as a way of refraining from a reliance on logocentric texts and of giving more practical opportunities for students to exercise a metalanguage. The results also align with the broader pedagogical shift in the field of ESP, deviating from rigid grammatical structures and rote memorization in favour of more experiential and multimodal learning approaches (Christoforou & Efthimiou, 2023). The role of 21st-century ESP students as future graduates needs to evolve, from passive recipients of linguistic information to active *designers* of course content. Although language remains an integral part of literacy education, integrating multimodal literacy ensures that students develop the diverse literacy skills required to participate in the digital age.

Second, these findings have broader pedagogical implications beyond ESP or engineering-specific contexts, inviting educators to rethink the use of Gen AI tools

in their practice, not solely as content creator tools, but as collaborative meaning-making resources within pedagogically designed learning environments. Finally, the study promotes more equitable ways of using Gen AI tools as meaning-making resources, as students constructed their own individualized meanings rather than passively relying on AI-generated outputs. This addressed the "cheating" concern, a common worry among instructors, since students remained critical of the Gen AI-generated results (Kalantzis & Cope, 2025a).

Conclusion

The study highlights significant pedagogical implications for, both, teachers and students in the context of ESP education. With regard to teachers, the study can have implications for teacher professional development in contexts where AI-assisted multimodal pedagogies are being adopted. As educators begin to incorporate GenAI tools into their teaching, training may be needed to help them guide students in exploring the multimodal affordances of these technologies. This can support a shift in how students engage with course content since teachers as *designers* of the students' learning experiences, especially through multimodal compositions, can help increase students' agency by involving them in the learning process and helping them reflect on the knowledge they have learnt. Through that process, students also become more socially responsible and demonstrate skills, such as creativity, critical thinking and semiotic awareness (Lim & Tan-Chia, 2022). Moreover, the study reinforces the need for ESP curricula to integrate the grammar of transposition and metalanguage as a foundational pedagogical framework that will help redesign academic literacy in digital contexts. Specifically, the students' use of transpositional strategies to shift meaning across forms (e.g. from verbal to visual or audio), highlights the importance of the grammar of transposition in guiding multimodal composition. Furthermore, the structured use of metalanguage enables students to reflect on their *design* choices, making the learning process more engaging. Although the current study expanded our understanding of the usefulness of Gen AI-assisted transpositional meaning *designs* through the creation of multimodal artefacts, it is not without its limitations. First, the small sample size and the homogeneity of the participants may impact the generalizability of the findings. Even though the composition of the all-female participant group reflected the demographics of the class and was not a deliberate choice, it may have influenced certain aspects of the data such as the multimodal design choices, collaboration dynamics between the groups, or patterns of commitment to the project. Future research could focus on how gender diversity might shape design choices in different educational contexts. Moreover, while transpositional grammar has shown promise in enhancing ESP students' multimodal literacy, further research is required to understand its effectiveness, in particular in the context of Gen AI usage, ensuring that ESP pedagogy remains inclusive and adaptive to the digitally evolving global context.

Acknowledgments

Sincere appreciation is extended to the ESP students of *Chemical Engineering I* for their dedication and enthusiasm throughout this project.

Author contributions

CRediT: **Maria Christoforou:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Maria Christoforou is an ESP instructor at the Cyprus University of Technology, where she has been designing and teaching English for Specific Purposes courses since 2012. She has been involved in English language teaching since 2002 and is currently pursuing a PhD exploring the pedagogical application of Virtual Reality in multimodal EFL learning. Her research interests include VR-assisted language learning, immersive technologies, and multimodality.

ORCID

Maria Christoforou http://orcid.org/0000-0001-7598-6159

Consent to participate

Informed consent to participate was obtained from all individual participants included in the study.

Consent for publication

Participants signed consent forms permitting the researcher to analyse their data.

Data availability statement

The datasets produced and examined in this study can be obtained from the corresponding author upon reasonable request.

References

Christoforou, M. (2022). University students' transformation of meanings within an ESP digital context. *Professional & Academic English*, *29*(2), 23–45.

Christoforou, M., & Efthimiou, F. (2023). Introducing *Dreams of Dali* in a tertiary education ESP course: Technological and pedagogical implementations. In P. Zaphiris & A. Ioannou (Eds.), *Learning and collaboration technologies. HCII 2023. Lecture notes in computer science* (Vol. 14041, pp. 53–65). Springer. https://doi.org/10.1007/978-3-031-34550-0_4

Cope, B., & Kalantzis, M. (Eds.). (2000). *Multiliteracies: Literacy learning and the design of social futures*. Routledge.

Cope, B., & Kalantzis, M. (2009). "Multiliteracies": New literacies, New learning. *Pedagogies: An International Journal*, *4*(3), 164–195. https://doi.org/10.1080/15544800903076044

Cope, B., & Kalantzis, M. (2020). *Making sense: Reference, agency, and structure in a grammar of multimodal meaning*. Cambridge University Press. https://doi.org/10.1017/9781316459645

Cope, B., & Kalantzis, M. (2023). Towards education justice: Multiliteracies revisited. In G. C. Zapata, M. Kalantzis, & B. Cope (Eds.), *Multiliteracies in international educational contexts: Towards education justice* (pp. 1–26). Routledge.

Cope, B., & Kalantzis, M. (2024). A multimodal grammar of artificial intelligence: Measuring the gains and losses in generative AI. *Multimodality & Society*, *4*(2), 123–152. https://doi.org/10.1177/26349795231221699

Creswell, J. (2015). *Educational research. Planning, conducting, and evaluating quantitative and qualitative research* (5th ed.). Pearson Education Inc.

Fernández-García, A. (2024). Musical input and a multiliteracies approach to facilitate intercultural EFL learning. *International Journal of English Studies*, *24*(2), 1–19. https://doi.org/10.6018/ijes.554961

Godwin-Jones, R. (2024). Generative AI, pragmatics, and authenticity in second language learning. *arXiv:2410.14395*.

Hafner, C. A. (2025). Multimodality and ESP research. In S. Starfield & C. A. Hafner (Eds.), *The handbook of English for specific purposes* (2nd ed., pp. 619–639). John Wiley & Sons, Inc.

Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Edward Arnold.

Jewitt, C. (2005). *Technology, literacy, learning: A multimodal approach*. Routledge. https://doi.org/10.4324/9780203964101

Kalantzis, M., & Cope, B. (2022). After language: A grammar of multiform transposition. In C. Lótge (Ed.), *Foreign language learning in the digital age* (pp. 34–64). Routledge.

Kalantzis, M., & Cope, B. (2025a). Literacy in the time of artificial intelligence. *Reading Research Quarterly*, *60*(1), e591. https://doi.org/10.1002/rrq.591

Kalantzis, M., & Cope, B. (2025b). Multiliteracies since social media and artificial intelligence. *Harvard Educational Review*, *95*(1), 135–151. https://doi.org/10.17763/1943-5045-95.1.135

Karimi, H., Sañosa, D. J., Hernandez Rios, K., Tran, P., Chun, D. M., Wang, R., & Arya, D. J. (2023). Building a city in the sky. *CALICO Journal*, *40*(1), 24–44. https://doi.org/10.1558/cj.22838

Kern, R. (2000). *Literacy and language teaching* (1st ed.). Oxford University Press.

Kress, G., & Selander, S. (2012). Multimodal design, learning and cultures of recognition. *The Internet and Higher Education*, *15*(4), 265–268. https://doi.org/10.1016/j.iheduc.2011.12.003

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.

Lee, J. C. P. (2024). Co-speech gestures in enhancing engineering proposal presentations: A multimodal analysis of ESP spoken discourse. *ESP Today*, *12*(2), 319–348. https://doi.org/10.18485/esptoday.2024.12.2.6

Lee, S.-Y., Lo, Y.-H. G., & Chin, T.-C. (2019). Practicing multiliteracies to enhance EFL learners' meaning making process and language development: A multimodal problem-based approach. *Computer Assisted Language Learning*, *34*(1–2), 66–91. https://doi.org/10.1080/09588221.2019.1614959

Lim, F. V. (2020). *Designing learning with embodied teaching: Perspectives from multimodality* (1st ed.). Routledge. https://doi.org/10.4324/9780429353178

Lim, F. V., Cope, B., & Kalantzis, M. (2022). A metalanguage for learning: Rebalancing the cognitive with the socio-material. *Frontiers in Communication*, *7*, 830613. https://doi.org/10.3389/fcomm.2022.830613

Lim, F. V., & Tan-Chia, L. (2022). *Designing learning for multimodal literacy: Teaching viewing and representing* (1st ed.). Routledge. https://doi.org/10.4324/9781003258513

Mills, K. A. (2011). *The multiliteracies classroom*. Multilingual Matters. https://www.multilingual-matters.com/page/detail/The-Multiliteracies-Classroom/?k=9781847693181

Mills, K. A., Unsworth, L., & Scholes, L. (2023). *Literacy for digital futures: Mind, body, text* (1st ed.). Routledge. https://doi.org/10.4324/9781003137368

New London Group, T., & New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, *66*(1), 60–92. https://doi.org/10.17763/haer.66.1.17370n67v22j160u

Paltridge, B. (2013). Genre and English for specific purposes. In B. Paltridge & S. Starfield (Eds.), *The handbook of English for specific purposes* (pp. 347–366). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118339855.ch18

Parkinson J.(2019).Multimodal student texts: Implications for ESP. In K. Hyland, L.L.C. Wong,(Eds.), *Specialised English: New directions in ESP and EAP research and practice* (1sted., pp.149–162). Routledge. https://doi.org/10.4324/9780429492082

Plastina, A. F. (2013). Multimodality in English for specific purposes: Reconceptualizing meaning-making practices. *LFE Revista de Lenguas para Fines Específicos*, *19*, 372–396.

Rahmanu, I. W. E. D., & Molnár, G. (2024). Multimodal immersion in English language learning in higher education: A systematic review. *Heliyon*, *10*(19), e38357. https://doi.org/10.1016/j.heliyon.2024.e38357

Reinhardt, J., & Thorne, S. (2011). Beyond comparisons: Frameworks for developing digital L2 literacies. In N. Arnold & L. Ducate (Eds.), *Present and Future Promises of CALL: From Theory and Research to New Directions in Language Teaching*, (Vol. *2*, pp. 257–280). CALICO.

Robson, C., & McCartan, K. (2016). *Real world research: A resource for users of social research methods in applied settings*. Wiley.

Saldaña, J. M. (2021). *The coding manual for qualitative researchers* (4th ed.). SAGE Publications.

Silitonga, L. M., Wiyaka, W., & Prastikawati, E. F. (2024). Boosting students' ESP vocabulary by utilizing AI chatbot. *Eternal (English Teaching Journal)*, *15*(2), 275–283. https://doi.org/10.26877/eternal.v15i2.605

Thrasher, T., Christoforou, M., & Ijiri, A. (2024). Virtual reality-mediated language learning: A case study of immersive learning in the Metaverse. In R. Vurdien & W. Chambers (Eds.), *Technology-mediated language learning and teaching*. IGI Global (pp. 25–57). IGI Global. https://doi.org/10.4018/979-8-3693-2687-9

Tzirides, A., Galla, M., Cope, B., & Kalantzis, M. (2025). Thinking through AI: Advancing cognitive and collaborative research for AI in education. https://doi.org/10.35542/osf.io/s8hqe

Waly, M. M., & Zakiyyah, A. M. (2024). Artificial intelligence assistance in English for communication competence of English for specific purposes. *INSANIA: Jurnal Pemikiran Alternatif Kependidikan*, *29* (2), 236–256. https://doi.org/10.24090/insania.v29i2.11140