INNOVATION IN LANGUAGE LEARNING International Conference

Creation and Implementation of AI-Generated Educational Videos for an Information Technology Glossary in English for Specific Purposes Courses

Andreja Kovačić[1], Goran Bubaš[2]

[1]University of Zagreb Faculty of Organization and Informatics, Croatia
[2] University of Zagreb Faculty of Organization and Informatics, Croatia

Abstract

Artificial intelligence tools for generating videos have recently become relatively inexpensive for creating brief, high-quality videos that contain speech and sound. Such videos can depict specific and complex scenes, including those in real-world professional settings. For English for Specific Purposes (ESP) instruction, these affordances create a novel opportunity for custom-made educational multimodal materials or modules adapted to specialized topics, vocabulary, and lesson segments. This paper provides a preliminary report on the creation and implementation of a glossary for two ESP courses related to Information Technology (IT) and Software Engineering. In creating the glossary, the video-generating feature of the Google Gemini AI tool was utilized. The main idea was to represent key technical terms extracted from course teaching materials and exams. Each generated glossary-related video contained three to six technical terms contextualized in a brief dialogue between actors. The videos were made available on the course webpages within a learning management system (LMS) and embedded as lessons based on microlearning pedagogy, together with brief textual descriptions of the included technical terms. Such an approach to integrating AI-generated videos provided students and the instructor with: (a) a means of learning or consolidating technical terms; (b) a potentially appealing online learning resource that students could use to expand their lexical knowledge and improve their comprehension of course material. In addition, the glossary was conceived as an informal tool for initial assessment or self-assessment of students' familiarity with technical jargon, since it was implemented at the beginning of the ESP courses in the first semester and study year. Finally, owing to the quality pronunciation of the technical terms in the videos—all of them pronounced in American English—students were also exposed to target ESP vocabulary in an audio (multimedia) format and given an opportunity to check and improve their pronunciation of selected technical terms. A brief empirical evaluation of the AI-generated video glossary is also provided.

Keywords: *English for Specific Purposes, GenAI Tools Integration, Video Glossary, Technical Jargon, Pronunciation*

1. Introduction

According to Ethnologue [1] and other statistical sources, the English language is predominant worldwide as a spoken language, ranking first with approximately 1.5 billion speakers, whereas, as a native language, it ranks only third with less than 400 million speakers in 2025, out of a world population of 7.6 billion. It is estimated that there are more than 20,000 higher education institutions (HEIs) in 190 countries and territories (see: [2], [3]). As a cautious estimate, based on the search of the World Higher Education Database [2], there may be thousands of HEIs worldwide that offer English for Special Purposes (ESP) or related courses in non-native English-speaking countries with a particular focus on information technology or computer science. Creating an online video glossary with the use of large language models for such a course was the main topic of the study presented in this paper.

Looking more broadly, a recent review study [4] revealed that ESP-related research is predominantly associated with courses offered in the following discipline categories: Social Sciences, Health Sciences, Applied Sciences and Technology, and Humanities. Another earlier analysis [5, pp. 12-14] categorized the applications of ESP and EAP (English for Academic Purposes) courses in the following most frequent categories: STEM, Medicine/Nursing, Law, Business, Tourism/Hospitality, Social Sciences, and Humanities.

The use of multimodality (combining text with images, audio, and video) has the potential to contribute to learners' language proficiency in ESP and EAP courses, including their English vocabulary [6]. Similar findings have been reported in relation to more advanced technologies like virtual reality (VR), which can also enhance key language competencies such as speaking, writing, and vocabulary acquisition

[7]. In fact, one study demonstrated that combining AI and extended reality (XR) could specifically benefit vocabulary acquisition [8].
Empirical findings reveal that *AI-ready pedagogies* have the potential to improve EFL learning outcomes, including writing accuracy [9]. In the study presented in this paper, a multimodal approach involving the use of AI-generated videos accompanied with text was utilized to design an online glossary of IT- and computer science-related terms for an ESP higher education course.

1.1. Use of Video in Education

Adding video content to existing teaching materials can enhance learning [10], while the use of multimedia tools can increase student engagement and motivation, as well as result in greater performance when used with proper pedagogy [11]. Numerous video characteristics may influence the use of video for teaching and learning [12]. The identified benefits of using AI to generate instructional videos include effectiveness, scalability, and personalization [13].
Since 2020, there has been a growing number of AI-based video generation tools (see: [14], [15]) and an increasing need for appropriate evaluation frameworks, including human-based evaluation. By 2026, there had also been rapid advances of short-form AI-generating video platforms that are primarily based on text-to-video (T2V), image-to-video (I2V), and video-to-video (V2V) generation [16], with a possible effect of overexposure to AI-based videos that may result in perceived AI-video content overload [17].
The study that is presented in this paper focuses on AI-generated short-form videos for the design of a multimodal (video, audio, and text) glossary in an ICT- and computer science-related ESP course with a focus on production effectiveness and minimizing technology related costs. Another objective was that the predominantly AI-generated multimodal video glossary would be evaluated by students as human users.

1.2. AI-generated Videos in ESP/EAP Courses

A recent review study has confirmed the effectiveness of AI in language skills learning and teaching, with the strongest effects on vocabulary learning [18], and AI working best as a supplement to, rather than a replacement for traditional teaching. Likewise, video-based instruction can significantly support vocabulary acquisition and retention in ESP courses [19]. Recent empirical studies indicate that AI-generated educational videos, compared to human-created educational videos, may produce similar positive effects on learning outcomes [20], [21], and may sometimes outperform traditional instruction methods, for instance, when applied in the specific field of medical education [22].

1.3. AI-Generated Videos for Vocabulary Acquisition

While research on the use of AI-generated videos for vocabulary acquisition in English as a Foreign Language (EFL) and ESP/EAP courses is still relatively scarce (see, for instance, [21]), one empirical study [23] indicated not only predominantly positive learner perceptions of such videos by *higher education* students, but also that “visual and auditory elements supported their understanding of professional vocabulary in meaningful situations”. It should also be noted that a similar positive effect of the use of AI-generated videos on English vocabulary learning was observed in a small-scale study involving Chinese *junior primary school* students [24]. Positive findings regarding vocabulary learning were also revealed in a study where activities involving learner-generated videos outperformed viewing of videos in a control group [25], as well as in a study where the students in the experimental group used an AI tool to generate cartoon movies, whereas students in the control group utilized the traditional non-AI based presentation–practice–production approach [26]. Finally, one study provided indirect evidence that the effectiveness of AI-generated instructional videos may depend on the integration of textual and visual information [27].

2. Main Goal and Research Questions of the Study

The study research focus presented in this paper is placed within the areas of second language learning and ESP/EAP, the use of videos in glossary creation, the use of AI-generated scenarios and videos, as well as AI-generated textual definitions of glossary words. In addition, the students’ evaluations of the AI-generated video glossary are examined.

The **main goal** of this study was to develop a time-efficient, low-cost model for creating an AI-generated video glossary for ESP and to evaluate students' perceptions of its effectiveness, usefulness, and motivational value for learning specialized vocabulary.
Based on the previously defined *main goal*, the following ***research questions*** were formulated:

RQ1: What are the cost and time requirements for developing a fully AI-generated video glossary in a specialized ESP course?

RQ2: How do students perceive the effectiveness of the AI-generated video glossary in acquiring and remembering specialized ESP vocabulary?

RQ3: How do students perceive the contribution of brief, contextualized, AI-generated videos with supplementary text to their understanding of technical terminology?

3. Model for the Time-effective and Low-Cost Development of an AI-Generated Video Glossary

In September 2026, Google's AI tool Gemini, with a basic paid subscription plan, enabled users to create up to seven realistic videos per day with spoken audio (high quality, natively generated videos with lip-sync or facial synchronization from typed English dialogue), each lasting 10 seconds. The quality of those videos was sufficiently high to make them difficult to distinguish from actual video recordings of speaking human actors. Therefore, if video-creation by Google Gemini takes place over three to four weeks, with a rate of seven videos created per day, this could provide a low-cost means to generate more than 150 videos for an ESP/EAP course. Even assuming a pessimistic rate of 50% or more of inadequate AI-generated videos it would still enable a considerable number of videos to be produced at the cost of the basic subscription within a reasonable period. Creating a realistic AI-generated video featuring human actors speaking requires a scenario for each video scene (a detailed prompt or set of instruction for text-to-video AI generation). Writing these scenarios may involve considerable time and effort on the part of the ESP/EAP teacher and can be automated by delegating this task to an LLM like ChatGPT or Gemini. Finally, for the video glossary, a list of technical terms has to be created that will be used in AI-generated videos. This task can also be delegated to an LLM providing either a description of the ESP/EAP course content or by uploading learning resources, tests, or both to the LLM. Finally, an LLM can not only create lists of technical terms to be included in educational videos and scenarios that utilize those words, but also provide definitions both in English and the students' primary language (i.e., translations). Of course, the scenarios, the AI-generated videos, the definitions of technical terms in English, and the translations of the terms and definitions to the students' first language (L1) should be monitored, evaluated, adapted, and corrected by the ESP/EAP teacher and/or language expert.
Basically, the specific work on the creation of a video glossary for the ESP courses evaluated in this study was carried out in the following steps:
1. The latest version of ChatGPT available in September 2026 (with a basic subscription) was asked to read a RAR archive containing 12 Microsoft Word files with course material and midterm exams, written in English language, for an ESP course related to information technology and software engineering.
2. ChatGPT was then asked to make a list of most frequent words related to software engineering, computer science, and technology appearing in those documents by examining the documents individually and counting the occurrence of each technical word or concept in all of the documents.
3. After creating a list of more than 80 unique technical terms that appeared in the ESP course related documents (learning material and exams) and their frequency of occurrence count, ChatGPT was asked to rank those terms according to their importance for teaching students of English for Special Purposes (ESP) courses enrolled in the first year of an undergraduate software engineering program at a Croatian university.
4. Finally, in the continuation of the same dialogue, ChatGPT was asked to create four exemplary scenarios for AI-generated videos using the following instruction: "I can create short videos of up to 15 seconds in length for free. The videos I intend to create will include (a) a description of the person speaking, (b) a description of the context or setting, such as a Zoom conversation or a teacher speaking in front of a classroom, and (c) where appropriate, a description of the listener and their response. The video-generation tool is excellent at creating realistic visual environments, facial expressions, and synchronizing them with spoken text. It also automatically incorporates other visual and contextual details that may not be explicitly mentioned in the prompt. Using the list of the most important technical terms, create four descriptions (prompts) for generating four different videos. Each prompt should include (1) the context and setting, including a description of the scene and the physical surroundings of the speaker(s), (2) a description of the speaker and listener, including their appearance, roles, and expressions, (3) the speaker's dialogue, with an engaging, natural-sounding script that incorporates selected technical terms and fits within a 15-second video, (4) where appropriate, a brief and natural reply from the listener, and (5) a variety of terminology, avoiding repetition of the selected technical terms across the four

prompts as much as possible. Make the dialogues engaging, informative, and appealing to first-year university students. The aim is to introduce technical terminology in an interesting and memorable way while placing it in realistic academic or professional situations. Ensure that the spoken dialogue has a natural conversational pace and fits comfortably within the 15-second time limit."

5. The appropriateness and effectiveness of the first scenarios generated by ChatGPT were tested using the Google Gemini's video-generation functionality. The created videos were each 10 seconds long and most of them were considered adequate ("ready to use") without alteration for inclusion in the final video glossary.

6. ChatGPT was repeatedly asked to create new scenarios while avoiding the use of technical terms that were included in the previous ones. Most scenarios included 3-6 technical terms with some exceptions (+/- 1-2 terms). Some additional technical terms considered important by ChatGPT for the ESP course were added to the initial list (regarding IT and software engineering studies) and the final list contained more than 100 different technical terms (ranging from "bit" and "RAM" to "software development" and "system engineer"). Occasionally, it was necessary to adapt the scenarios created by ChatGPT for diversity purpose and to make the videos slightly more interesting, but most of the scenarios initially created by ChatGPT were used unchanged. This use of ChatGPT-generated scenarios substantially reduced the time needed to create AI-generated videos. A "screenshot" of a sample scenario and an image of the resulting video are presented in *Fig. 1*.

7. In a sequence of days Gemini was used to create up to 7 videos per day and a total of 36 final versions of AI-generated videos were produced in less than 3 weeks. Fewer than 50% of the initial videos created by Gemini required correction before being included in the video glossary. The final version of the video glossary included 30 selected videos that illustrated the use of more than 120 different technical terms in context.

8. To create the video glossary in a Moodle learning management system (LMS) the "Quiz" functionality was used. First, a question bank was created, and then, broadly in order from simpler / basic to more advanced / specialized technical terms, the questions were placed in the final version of the Moodle quiz. Each quiz item was presented on a separate page and contained (from top to bottom) a written transcript of the spoken dialogue, an embedded 10-second video, a list and textual definitions of technical terms in the video in English, and their translations into Croatian, all in form of a checkbox question with technical terms and their definitions as optional answer choices. Example of an AI-generated video glossary item is presented in *Fig. 2*.

9. For evaluation purpose, each quiz item containing a video and set of technical terms was supplemented by two Likert-type questions regarding its perceived (1) usefulness, and (2) interestingness.

Having in mind the *effectiveness* of the previously described model of AI-generated video-glossary development for this specific ESP course, it is worth emphasizing that the total amount of work required to complete the previously described stages, including the production of 30 AI-generated videos embedded in Moodle quiz items, AI-generated descriptions and translations of 120 technical terms, and the creation of the Moodle quiz, was *less than 50 hours*. However, it must be noted that the second author of this paper, who performed the technical work using ChatGPT and Gemini LLMs, had considerable experience with LLM prompting techniques.

4. Methodology for the Evaluation of the Video Glossary

The data collected for the evaluation of the video glossary in the ESP course included the following: (1) the frequency with which students marked each technical term in the Moodle quiz items as the ones they had learned or became more familiar with through their use of the video glossary, but had not previously known well in English; (2) the evaluation of each quiz item in terms of usefulness and interestingness of the individual video and its accompanying text using a five-point Likert-type scale; (3) responses to a voluntary and anonymous online survey using Google Forms containing items related to the overall evaluation of the features of the video glossary. However, for the sake of brevity, only selected data from the online survey will be presented in the remainder of this paper.

4.1. Instruments

The instruments used for the empirical part of this study were: (a) an AI-generated video glossary implemented using the Quiz functionality in the Moodle LLM with included evaluation questions of each glossary item and its technical terminology; (b) an online survey created using Google Forms.

4.2. Respondents

The respondents (participants) in the following brief empirical research report were 81 first-year undergraduate university students enrolled in a software engineering study at the University of Zagreb, Croatia, who were attending an ESP course and competed the voluntary and anonymous online Google Forms survey (out of 121 students who were enrolled in the course and 91 who started their Moodle quiz with the AI-generated video glossary). The respondents were 80.2% of male and 19.8% of female gender. In the State Graduation Exam for English language 75.3% of the respondents received a grade 5 (excellent) or 4 (very good), and only 19.7% received a lower grade. Also, their last grade in English language course in secondary school was in 91.3% of cases also 5 (excellent) or 4 (very good). Those findings suggest, on the average, a relatively high initial level of general English proficiency of most of the participants in the study. Finally, 81.3% of the respondents confirmed that they had seriously or very seriously responded to survey items, while 92.3% stated that they had answered sincerely or very sincerely.

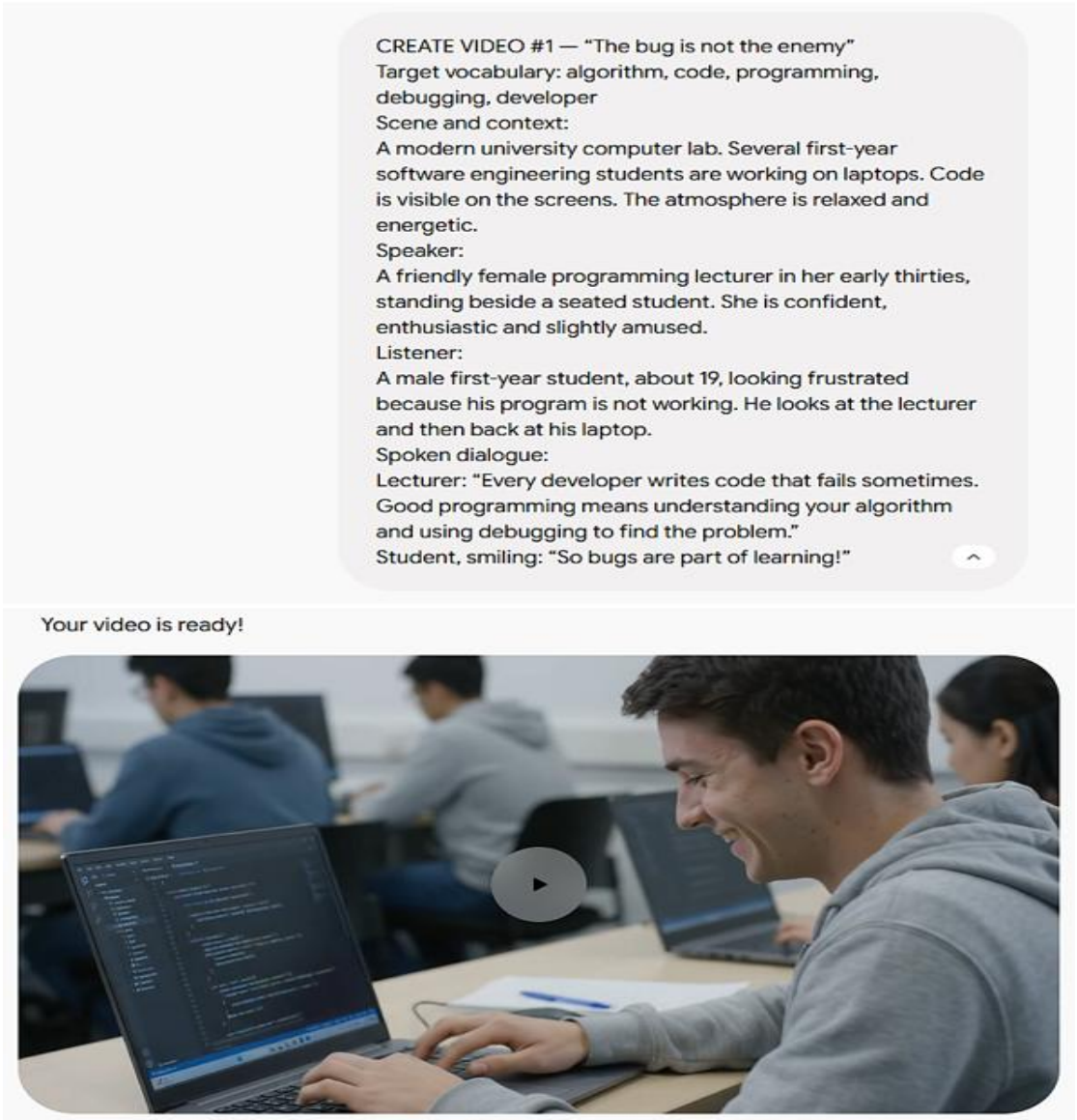


Fig. 1. Example of a ChatGPT-generated scenario and video output produced by Google Gemini

4.3. Procedure

After it was created (as described in Section 5. "Model for the Low-Cost Development of an AI-Generated Video Glossary"), the AI-generated video glossary was made available to students enrolled in the ESP courses at the beginning of the winter semester of their first year of undergraduate university study. Both

the use of the AI-generated video glossary and completion of the anonymous online survey were voluntary. The students had access to the video glossary and the online survey for 10 days. They were informed that most of the technical terms in the video glossary might appear in the questions of their two midterm exams. For the performing the previously outlined activities students were awarded a small amount of course credit. The study was approved by the institution's Ethics Committee.

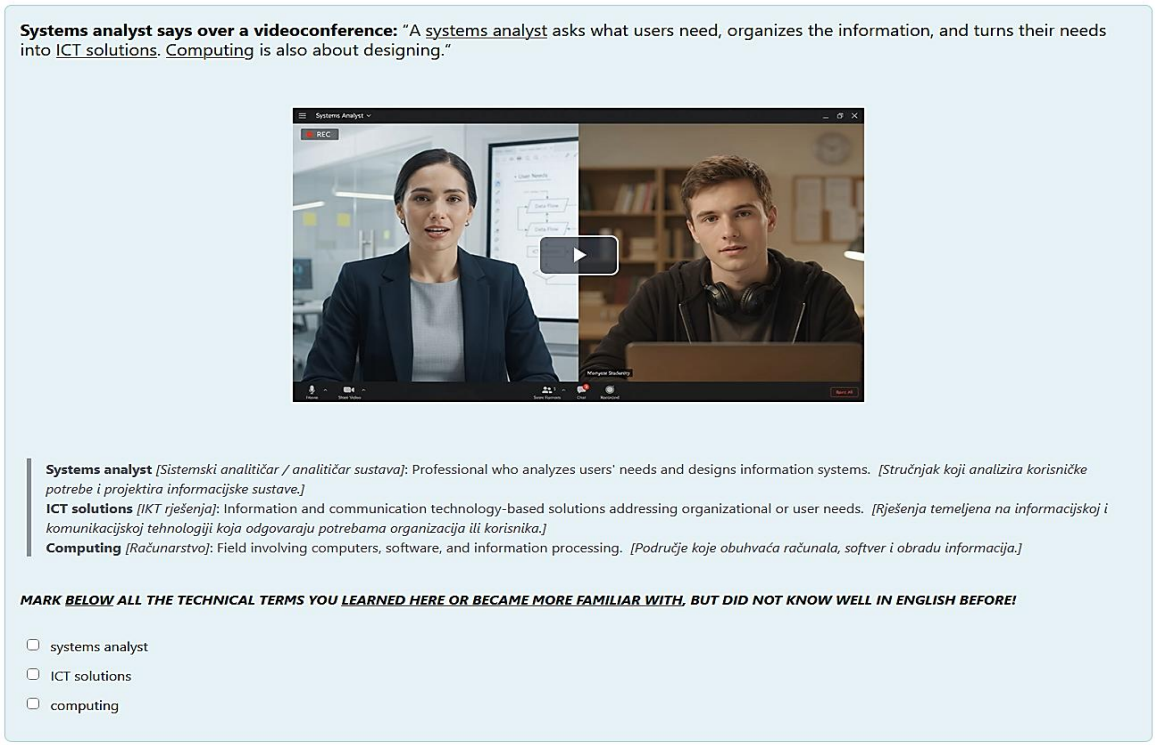


Fig. 2. Example of an AI-generated video glossary item featuring a transcript of the spoken dialogue, an embedded 10-second video, and definitions of the technical terms used

5. Results

This section briefly presents the main results of the anonymous and voluntary online survey (N = 81) focusing on the students' evaluations and perceptions of the AI-generated video glossary.

Regarding the survey item "Using the video glossary helped me better understand and learn certain terms related to informatics and computing", the most common responses were "Mostly agree" (44.4%) and "Strongly agree" (27.2%) bringing the total of positive responses to 71.6%. Also, the positive responses to the item "Do you think that combining textual content and videos helped you better remember technical terms you previously knew poorly or did not know?" were "Yes, but only to some extent" (38.3%) and "Yes, very much" (21.0%), with a total of 59.3% positive responses. Finally, the positive responses to the survey item "Does it help that thematically related words, i.e. technical terms that can appear together in speech and IT literature, were mentioned together in the videos and accompanying text?" were "Mostly yes" (58.0%) and "Yes, very much" (17.3%) and their total was 75.3%. Therefore, grouping thematically related words in AI-generated videos and supplementary text, combining technical terminology and videos to improve retention of technical terms, and the overall perception that the AI-generated video glossary was helpful for understanding and learning technical terminology were its highest rated aspects.

The percentages of positive responses to the survey question "How much did the examples of English technical words in the videos and accompanying texts help you better understand the specialized vocabulary of informatics and computing?" were 34.6% for "Much" and 7.4% for "Very much", bringing the total of positive responses to 41%. Similarly, the positive responses to the survey question "Did the use of realistic videos help you gain a deeper/fuller understanding of certain previously little-known technical terms and the situations in which they are used?" were "Much" (21.0%) and "Very much" (7.4%), having the total of positive responses of 28.4%. As a final point, the positive responses to the survey question "Compared with other learning methods, how do you rate the use of short (AI-generated) videos combined with textual explanations

of technical terms?" were "Moderately positive" (30.9%) and "Very positive" (9.9%), totaling 40.8%. The predominant response for the previous three survey items was *neutral* (ranging from 35.8% to 38.3%) which suggested that the AI-generated videos and text for the video glossary achieved, on average, both positive and mild/neutral perceptions and evaluations.
Other survey results can be briefly summarized as follows: *textual content* (accompanying the videos) for learning English technical terminology in informatics and computing was, on the average, better evaluated than the AI-generated videos; the duration of the videos was perceived as "Optimal" by 71.5% of the respondents; 75.3% of all respondents reacted positively (by selecting "Much" or "Very much") to the survey question "How relevant was the content of the videos and accompanying text to your study program at the faculty?".

6. Conclusion

The theoretical rationale for this approach within ESP teaching and learning for implementing such a pedagogical approach to vocabulary instruction of technical terms through a multimodal online IT glossary is grounded in the following key principles:
(1) It represents a well-established combination of language-centered and learning-centered approaches to ESP course design [28] which exposes students to the language (vocabulary) that they will need to use in future professional situations, while enabling them to assess the current state of their existing familiarity with the lexical material provided in the glossary.
(2) Incorporating a collection of more than 100 lexical items (acronyms, words, and phrases) in a single resource at the beginning of the course that provides students with a broader view of the range of terms relevant to their field of study. In the AI-generated video glossary these terms were presented in form of short scenarios (dialogues) in which related terms are presented in a specific context. This approach is based on the notion that placing terms belonging to the same conceptual domain in the same context helps activate students' frame (background) knowledge [29] and "relate lexical units into cognitive structures" [30, p. 772].
(3) The meaningful integration of video, audio, and text in the video glossary is aimed to facilitate the processing and comprehension of lexical material. This feature of the glossary draws on Mayer's multimedia learning theory [31].
Overall, the **main goal** of the study presented in this paper was achieved. Using only the basic subscriptions plans for ChatGPT and Gemini, a fully AI-generated video glossary comprising 30 videos and more than 120 technical terms, together with their definitions and Croatian translations was developed in fewer than three weeks and with less than 50 hours of work (***RQ1***), although some AI-generated scenarios and videos required correction and teacher oversight remained necessary. Regarding ***RQ2***, more than half of the students perceived the glossary as *helpful for understanding and learning technical terms*, almost 60% reported that *combining video and text helped them remember poorly known technical terms*, and three quarters positively valued the *grouping of thematically related terms*. Regarding ***RQ3***, students recognized the *contribution of the glossary to their understanding of technical terminology*, but they attributed this contribution more to the supplementary text than to the videos, whose contribution to deeper understanding received predominantly neutral ratings. These findings suggest that creating AI-generated video glossaries can be a rapid, low-cost, and relevant supplementary resource for ESP vocabulary learning, while future work should focus on improving the instructional value and motivational appeal of the videos and on measuring actual vocabulary acquisition and retention.

REFERENCES

[1] Eberhard D. M., Simons G. F., Robinson A. J. (eds.), *Ethnologue: Languages of the World*, 29th ed., Dallas, TX, SIL Global, 2026, https://www.ethnologue.com/.
[2] International Association of Universities (IAU), "About", World Higher Education Database (WHED), 2026, https://whed.net/About.html.
[3] UNIRANKS, "About WEBOMETRICS by UNIRANKS", Webometrics, 2026, https://www.webometrics.org/about-us.
[4] Yan H., "Trends in Empirical Research in English for Specific Purposes: A Systematic Review of SSCI-Indexed Journal Articles (2014–2023)", SAGE Open, vol. 15, no. 1, 2025, https://doi.org/10.1177/21582440251328460.
[5] Anthony L., *Introducing English for Specific Purposes*, London, Routledge, 2018, https://doi.org/10.4324/9781351031189.

[6] Rahmanu I. W. E. D., Molnár G., “Multimodal immersion in English language learning in higher education: A systematic review”, Heliyon, vol. 10, no. 19, 2024, article e38357, https://doi.org/10.1016/j.heliyon.2024.e38357.
[7] Rahmanu I. W. E. D., Molnár G., “Virtual reality in English language education: A systematic review of opportunities and challenges in higher education”, Acta Psychologica, vol. 264, 2026, article 106358, https://doi.org/10.1016/j.actpsy.2026.106358.
[8] Divekar R. R., Drozdal J., Chabot S., Zhou Y., Su H., Chen Y., Zhu H., Hendler J. A., Braasch J., “Foreign language acquisition via artificial intelligence and extended reality: Design and evaluation”, Computer Assisted Language Learning, vol. 35, no. 9, 2022, pp. 2332–2360, https://doi.org/10.1080/09588221.2021.1879162.
[9] Ekizer F. N., “Exploring the impact of artificial intelligence on English language teaching: A meta-analysis”, Acta Psychologica, vol. 260, 2025, article 105649, https://doi.org/10.1016/j.actpsy.2025.105649.
[10] Noetel M., Griffith S., Delaney O., Sanders T., Parker P., del Pozo Cruz B., Lonsdale C., “Video improves learning in higher education: A systematic review”, Review of Educational Research, vol. 91, no. 2, 2021, pp. 204–236, https://doi.org/10.3102/0034654321990713.
[11] Staneviciene E., Žekienė G., “The use of multimedia in the teaching and learning process of higher education: A systematic review”, Sustainability, vol. 17, no. 19, 2025, article 8859, https://doi.org/10.3390/su17198859.
[12] Navarrete E., Nehring A., Schanze S., Ewerth R., Hoppe A., “A closer look into recent video-based learning research: A comprehensive review of video characteristics, tools, technologies, and learning effectiveness”, International Journal of Artificial Intelligence in Education, vol. 35, no. 4, 2025, pp. 1631–1694, https://doi.org/10.1007/s40593-025-00481-x.
[13] Hai T. T., Mai D. T. T., Hanh N. V., “A rapid review of using AI-generated instructional videos in higher education”, Frontiers in Computer Science, vol. 7, 2026, article 1721093, https://doi.org/10.3389/fcomp.2025.1721093.
[14] Xie W., Hu A., Xie Q., Chen J., Wan R., Liu Y., “Bibliometric analysis and review of AI-based video generation: Research dynamics and application trends (2020–2025)”, Discover Computing, vol. 28, 2025, article 130, https://doi.org/10.1007/s10791-025-09628-9.
[15] Liu X., Xiang X., Li Z., Wang Y., Li Z., Liu Z., Zhang W., Ye W., Zhang J., “A survey of AI-generated video evaluation”, arXiv, article 2410.19884v2, 2026, https://doi.org/10.48550/arXiv.2410.19884.
[16] Wang R., “A comparative study on the practicality of AI video generation tools in short video creation”, Exploring Science Academic Conference Series, vol. 16, 2026, pp. 427–434, https://doi.org/10.70267/aitia.2026427434.
[17] Nguyen G. T. T., Nguyen H. T. M., “AI content, user fatigue, and churn on TikTok: Testing an integrated stressor model”, Human Behavior and Emerging Technologies, vol. 2026, no. 1, 2026, article 9977742, https://doi.org/10.1155/hbe2/9977742.
[18] Torres P. J., Kahveci Y. E., “Effectiveness of artificial intelligence (AI) in language teaching”, Computers and Education: Artificial Intelligence, vol. 9, 2025, article 100522, https://doi.org/10.1016/j.caeai.2025.100522.
[19] Liu X., Chang S., “A study of video-based ESP vocabulary learning and retention in a Chinese vocational university”, English Language & Literature Teaching, vol. 32, no. 2, 2026, pp. 51–71, https://doi.org/10.35828/etak.2026.32.2.51
[20] Netland T., von Dzengelevski O., Tesch K., Kwasnitschka D., “Comparing human-made and AI-generated teaching videos: An experimental study on learning effects”, Computers & Education, vol. 224, 2025, article 105164, https://doi.org/10.1016/j.compedu.2024.105164.
[21] Xu T., Liu Y., Jin Y., Qu Y., Bai J., Zhang W., Zhou Y., “From recorded to AI-generated instructional videos: A comparison of learning performance and experience”, British Journal of Educational Technology, vol. 56, 2025, pp. 1463–1487, https://doi.org/10.1111/bjet.13530.
[22] Impito P. F., “Artificial intelligence generated videos as supportive tools in medical education: A scoping review”, Journal of Medical Education and Curricular Development, vol. 13, 2026, https://doi.org/10.1177/23821205261437352.
[23] Alsofyani A., “The effect of AI-generated instructional videos on vocabulary acquisition among English as a foreign language learners”, EDULEARN26 Proceedings, Valencia, IATED Academy, 2026, article 2674, https://doi.org/10.21125/edulearn.2026.2674.
[24] Chan Y. H. C., Li P., “The impact of AI-generated videos on English vocabulary learning among newly arrived mainland Chinese children in Hong Kong”, 2026 7th International Conference on Information Technology and Education Technology (ITET), IEEE, 2026, pp. 74–79, https://doi.org/10.1109/ITET70052.2026.11651047.

[25] Xu C., Gu Y., “Creating vs. viewing multimodal videos: Effects on L2 vocabulary retention”, Proceedings of the 48th Annual Meeting of the Cognitive Science Society, Cognitive Science Society, 2026, https://escholarship.org/uc/cognitivesciencesociety/48/0.
[26] Behforouz B., Al Ghaithi A., “Visual language learning: Exploring the effects of artificial intelligence-generated animations on vocabulary and motivational growth”, Technology in Language Teaching & Learning, vol. 7, no. 4, 2025, article 103077, https://doi.org/10.29140/tltl.v7n4.103077.
[27] Pi Z., Liu C., Wang L., Yang J., Li X., “Cues facilitate foreign language vocabulary learning from instructional videos: Behavioral and neural evidence”, Language Teaching Research, vol. 30, no. 3, 2026, pp. 1005–1029, https://doi.org/10.1177/13621688231164724
[28] Hutchinson T., Waters A., *English for Specific Purposes: A Learning-Centred Approach*, Cambridge, Cambridge University Press, 1987.
[29] Fillmore C. J., “Frame semantics”, Linguistics in the Morning Calm, Seoul, Hanshin Publishing Co., 1982, pp. 111–137.
[30] Smirnova O., Rackevičienė S., Mockienė L., “Frame semantics methodology for teaching terminology of specialised domains”, Journal of Teaching English for Specific and Academic Purposes, vol. 9, no. 4, 2021, pp. 765–77, 2022, https://doi.org/10.22190/JTESAP2104765S.
[31] Mayer R. E., *Multimedia Learning*, 3rd ed., Cambridge, Cambridge University Press, 2021.