GENERATIVE AI-ASSISTED INTERSEMIOTIC TRANSLATION IN AUTOMOTIVE ENGINEERING ESP LEARNING: A CLASSROOM CASE STUDY

Phi Luong Van[1], Ngo Thi Lan Huong[1]

[1] University of Transport Technology, Hanoi, Vietnam.

Email: Phi Luong Van: vanpl@utt.edu.vn; Ngo Thi Lan Huong: huongntl@utt.edu.vn.

Abstract

This classroom case study examined the use of generative artificial intelligence (GenAI) in English for Specific Purposes (ESP) learning among 77 Automotive Engineering students. Students used ChatGPT and Canva AI to support intersemiotic translation by transforming technical content into infographics for oral presentations. Data were collected through classroom observations, analysis of student-generated learning artifacts, and a questionnaire survey. Drawing on multimodal semiotics and the Cognitive Theory of Multimedia Learning (CTML), the findings showed that students engaged in selecting, organizing, and verifying information and generally held positive views of GenAI use. However, some infographics contained inaccurate terminology and did not adequately represent technical relationships, highlighting the continued need for teacher guidance.

Keywords: *intersemiotic translation*; *multimodal semiotics; generative artificial intelligence; English for Specific Purposes (ESP); automotive engineering.*

Tóm tắt

Nghiên cứu tình huống tìm hiểu cách 77 sinh viên ngành Công nghệ kỹ thuật ô tô sử dụng trí tuệ nhân tạo tạo sinh (GenAI) trong học tiếng Anh chuyên ngành. Sinh viên dùng ChatGPT và Canva AI hỗ trợ dịch liên kí hiệu, chuyển nội dung kỹ thuật thành infographic phục vụ thuyết trình. Dữ liệu gồm quan sát lớp học, phân tích sản phẩm và khảo sát. Dựa trên kí hiệu học đa phương thức và lý thuyết nhận thức về học tập đa phương tiện, kết quả cho thấy sinh viên tham gia lựa chọn, tổ chức và kiểm chứng thông tin, đồng thời nhìn chung đánh giá tích cực sự hỗ trợ của GenAI. Tuy nhiên, một số infographic còn sai sót về thuật ngữ và cách thể hiện quan hệ kỹ thuật, cho thấy vẫn cần có sự hướng dẫn của giảng viên.

Từ khóa: *dịch liên kí hiệu; kí hiệu học đa phương thức; trí tuệ nhân tạo tạo sinh; tiếng Anh chuyên ngành; công nghệ kỹ thuật ô tô.*

1. Introduction

In English for Specific Purposes (ESP), automotive engineering students need to understand technical terminology and use it to explain component functions, relationships, and operating principles in English. Infographics and diagrams can help students present this information through text, images, and symbols. ChatGPT and Canva AI provide additional support for explaining terminology, selecting content, and preparing visual materials, although the accuracy of their outputs remains a concern.

Previous studies have examined conversational AI in language learning (Bibauw et al., 2019; Huang et al., 2022) and the use of ChatGPT in language education (Kohnke et al., 2023). Hafner and Ho (2026) examined multimodal composing with GenAI, particularly learner agency and the relationship between different modes. In Vietnam, Vo and Nguyen (2024), Pham and Le (2024), and Ho (2024) reported generally positive student perceptions of ChatGPT, while also identifying concerns about reliability and dependence on the tool. Similar

issues concerning information relevance and academic integrity were noted in the review by Pham and Cao (2025).

These studies have mainly examined students' perceptions and experiences of GenAI in language learning. Less is known about how ESP students use these tools to select technical information and represent it through text and images. In automotive engineering, this process requires particular care because diagrams may appear clear while containing inaccurate labels or misleading representations of technical relationships.

Jakobson's (1959) concept of intersemiotic translation concerns the interpretation of verbal content through images and other signs. When using ChatGPT to explain and summarize content and Canva AI to support infographic design, students need to select, check, and revise information before incorporating it into their products and presenting it in English. These activities relate to the processes of selection, organization, and integration in Mayer's (2009, 2021) Cognitive Theory of Multimedia Learning.

This study examined how students majoring in Automotive Engineering at the University of Transport Technology, Vietnam, used ChatGPT and Canva AI to develop infographics and present technical topics in English. Data from classroom observations, student-generated artifacts, and a post-project questionnaire were used to address three research questions:

RQ1. To what extent were content transformation, coordination of modes, information selection, and evaluation of GenAI outputs evident in the observed learning activities?

RQ2. How did students re-represent automotive terminology and technical content in infographics in terms of technical accuracy, text–image coordination, and information organization?

RQ3. What benefits and difficulties did students perceive when using ChatGPT and Canva AI to process content, design visual products, and support ESP learning?

2. Theoretical Framework

The study drew on two complementary frameworks to examine students' experiences of using GenAI in Automotive Engineering ESP learning: semiotics and translation, and the Cognitive Theory of Multimedia Learning.

2.1. Semiotics and Translation Framework

This framework provides a basis for analyzing the transformation of knowledge from text into visual representations. Jakobson (1959) defined intersemiotic translation as the interpretation of verbal signs through systems of nonverbal signs. In Automotive Engineering ESP learning, students' use of ChatGPT and Canva AI to transform terminology and technical principles into infographics and energy-flow diagrams can be viewed as intersemiotic translation, through which specialized content is re-represented visually.

This process can be examined further through Kress and van Leeuwen's (2006) grammar of visual design, drawing on the ideational, interpersonal, and textual metafunctions. Combining text, images, symbols, arrows, and spatial layout allows students to use the distinct meaning-making resources of each mode to clarify the structure and technical relationships of Automotive Engineering systems.

2.2. Cognitive Theory of Multimedia Learning

While the semiotics and translation framework explains how content is transformed into multimodal products, Mayer's (2009, 2021) Cognitive Theory of Multimedia Learning (CTML) provides an account of how learners receive and process information.

CTML rests on three assumptions: the cognitive system processes information through visual/pictorial and auditory/verbal channels (dual channels); each channel has limited processing capacity (limited capacity); and meaningful learning requires learners to select, organize, and integrate information actively (active processing).

In automotive engineering ESP learning, students work with technical terms, images, diagrams, and oral explanations. Selecting and organizing this information may support understanding, although the use of multiple modes does not necessarily lead to better learning.

2.3. Relationship between the Frameworks

The two frameworks serve different purposes in this study. Intersemiotic translation helps explain how technical content is represented across modes, while CTML provides a basis for examining how students select and organize information. Figure 1 brings these perspectives together. In practice, students may move back and forth between selecting content, designing visuals, and checking their work.

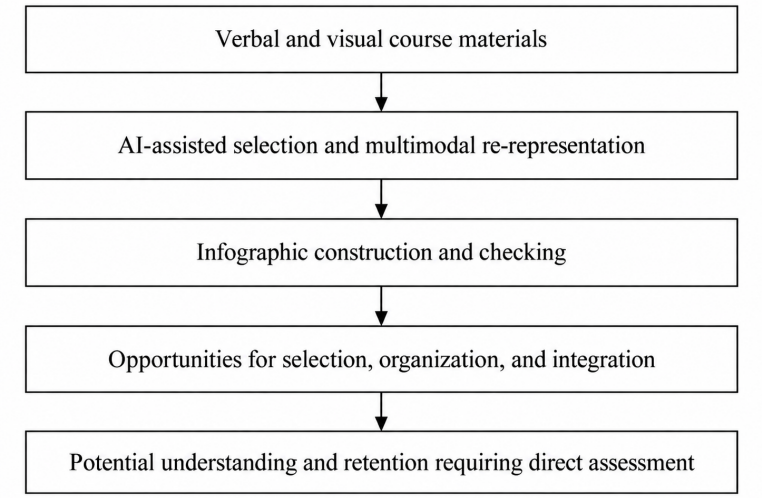


Figure 1. Proposed process of intersemiotic transformation and cognitive processing in GenAI-supported ESP learning

Figure 1 is a conceptual model, not an empirically validated causal pathway. Selection, organization, and integration are theoretical processes; the study documents observable task practices and product features. Understanding and retention remain potential outcomes requiring direct assessment.

3. Research Methods, Scope, and Participants

A classroom case study was conducted to examine how automotive engineering students used GenAI to transform verbal technical content into multimodal visual representations in ESP classes. It focused on ChatGPT and Canva AI for processing, explaining, and re-representing terminology, concepts, and principles related to automotive engineering through infographics, energy-flow diagrams, and materials for oral presentations.

3.1. Participants and Scope

Participants were 77 students majoring in Automotive Engineering, enrolled in two ESP classes at the University of Transport Technology, Vietnam. They had completed two general English courses before taking ESP. Both classes completed the ESP course in the second semester of the 2025–2026 academic year.

Students worked in groups, each presenting a technical topic in English using an infographic. Questionnaire data were collected from both entire classes after the learning tasks, while observations and artifact analysis focused on two groups across six units. The study was bounded by one course at one institution and examined GenAI practices, product features, and student experiences. It did not seek to assess overall English proficiency or compare AI platforms.

3.2. Research Design and Procedure

The study was embedded in classroom learning tasks, with the coursebook serving as the primary source for infographic content. The teacher provided prompts to guide students in using ChatGPT to identify terminology, explain technical concepts, and select relevant information. Students then used AI features in Canva for Education to create infographics, structural diagrams, or energy-flow diagrams from the selected content. This process involved both transforming verbal information into visual representations and reorganizing existing coursebook images and diagrams. Students checked AI outputs against the coursebook and revised their products before using them in group presentations in English.

3.3. Data Collection Methods and Instruments

Data were collected from three sources: classroom observations, learning artifacts, and a post-project questionnaire. All 77 students completed the questionnaire. Observations and artifact analysis focused on two groups across six units, yielding 12 observation sheets and 12 infographics.

The research team developed the observation protocol and questionnaire based on the semiotics and translation framework, the Cognitive Theory of Multimedia Learning, and the learning tasks. The observation protocol comprised 12 criteria across three categories: intersemiotic transformation, multimodal meaning-making, and active information processing and evaluation of GenAI outputs. Each criterion included specific behavioral indicators and was rated as follows: 0 = not observed; 1 = partially observed, when the behavior occurred occasionally or required substantial support; 2 = clearly observed, when it was demonstrated clearly, independently, and appropriately; and N/A = not applicable, when it could not reasonably be observed in the particular episode. The unit of recording was each criterion within a group observation episode.

Learning artifacts were collected to examine the accuracy of terminology and technical content, text–image coordination, representation of technical relationships, and information organization. The questionnaire was administered through Google Forms and included questions about class membership and prior GenAI experience, nine Likert-scale statements, and two open-ended questions. The five-point scale was labeled 1 = strongly disagree; 2 = disagree; 3 = undecided/neutral; 4 = agree; and 5 = strongly agree. Each Likert statement received 77 responses, while the open-ended questions on benefits and difficulties received 57 and 56 responses, respectively, in Vietnamese. The complete questionnaire and response data were retained in Google Forms for cross-checking and verification.

3.4. Data Analysis

Quantitative data were analyzed using descriptive statistics, including frequencies, percentages, means (M), and sample standard deviations (SD). Likert responses were coded from 1 to 5, and agreement percentages were calculated by combining ratings of 4 and 5. Qualitative analysis was informed by the semiotics and translation framework and the Cognitive Theory of Multimedia Learning. The same teacher taught both classes, while a second teacher participated in observation, evaluation, and data coding. Codes were organized into three categories: intersemiotic transformation, multimodal meaning-making, and active information processing and evaluation of GenAI outputs. Coding units comprised individual criteria within each group observation episode, meaningful elements and relationships within infographics, and meaningful statements in open-ended responses. A response could contain multiple coding units. The two teachers compared their coding results and discussed differences with reference to the assessment criteria and corresponding data to reach agreement on final interpretations. No inter-rater reliability coefficient was calculated or reported. Open-ended responses were analyzed in Vietnamese, their original language, and quotations were translated into English for reporting. Findings from observations, learning artifacts, and the questionnaire were triangulated to identify similarities and differences across the three data sources.

4. Results and Discussion

4.1. Classroom Observation Results

Table 1. Distribution of Classroom Observation Ratings by Criterion (N = 12)

Code	Observation criterion	0	1	2	N/A	Valid n
OBS1	Using ChatGPT to explain or simplify technical content	0	5	3	4	8
OBS2	Transforming technical content into visual representations	0	4	6	2	10
OBS3	Selecting images, symbols, and signs appropriate to technical concepts	0	6	4	2	10
OBS4	Representing technical relationships through arrows, sequencing, and spatial layout	0	8	3	1	11
OBS5	Combining text and images to complement meaning	0	3	6	3	9
OBS6	Organizing information into a coherent multimodal structure	0	6	5	1	11
OBS7	Linking English terminology to corresponding images	0	4	8	0	12
OBS8	Using multimodal products to support oral explanations	0	0	10	2	10
OBS9	Selecting and prioritizing relevant technical information	0	3	9	0	12
OBS10	Checking AI-generated information	0	1	10	1	11
OBS11	Identifying and correcting inaccurate or unsuitable outputs	0	4	6	2	10
OBS12	Demonstrating learner agency in GenAI use	1	3	7	1	11

Note. *Each criterion had 12 possible observation episodes (two groups × six instructional units). Ratings were coded as 0 = not observed, 1 = partially observed, 2 = clearly observed, and N/A = not applicable. Valid n = 12 − N/A. Percentages reported in the text were calculated using valid observations for each criterion.*

Two student groups were observed across six units, resulting in 12 observation sheets. Each observation covered 12 criteria related to intersemiotic translation, multimodal meaning-making, and students' processing and evaluation of GenAI-generated content. The distribution of ratings and the number of valid observations for each criterion are presented in Table 1. As some activities could not be observed in every session, the number of valid observations varied across criteria. Cases marked as N/A were excluded when calculating percentages.

Transforming Technical Content into Visual Representations

Within the intersemiotic translation criteria, transforming verbal technical information into visual representations (OBS2) was clearly observed in 60.0% of valid episodes. In contrast, the proportion of clearly observed behaviors was lower for selecting appropriate visual signs (OBS3, 40.0%) and representing technical relationships (OBS4, 27.3%).

The observations showed that transforming technical descriptions into visual materials was more evident than selecting appropriate visual signs or representing relationships between components. From the perspective of intersemiotic translation (Jakobson, 1959), the main concern is whether the technical meaning is preserved when information is presented in another mode.

Using ChatGPT to explain or simplify technical content (OBS1) reached level 2 in 3/8 valid episodes (37.5%) and level 1 in 5/8 (62.5%). Four episodes were recorded as N/A because the activity had taken place before observation. This result reflects the timing of data collection and cannot support a conclusion that students made limited use of ChatGPT during task preparation.

Meaning-Making through Text, Images, and Speech

Combining text and images to complement meaning (OBS5) reached level 2 in 6/9 valid episodes (66.7%) and level 1 in 3/9 (33.3%); no valid episode was rated at level 0. Organizing information into a coherent multimodal structure (OBS6), however, reached level 2 in 5/11 episodes (45.5%) and level 1 in 6/11 (54.5%). Students thus coordinated modes, but their organization into a clear, consistent structure was predominantly partial.

Linking English terminology to corresponding images (OBS7) reached level 2 in 8/12 episodes (66.7%) and level 1 in 4/12 (33.3%). Terms were placed near or connected to the relevant components and processes, helping identify their referents in technical contexts. The difference between OBS7 and OBS4 suggests that identifying components and explaining how they work together may require different forms of instructional support.

Using multimodal products to support oral explanations (OBS8) reached level 2 in all 10/10 valid episodes (100.0%). Students used images, labels, and arrows during presentations rather than simply reading the text on their products. From Kress and van Leeuwen's (2006) perspective, this can be interpreted as coordination of semiotic resources in meaning-making. However, the observed use of visuals does not directly establish the accuracy of explanations, audience comprehension, or individual members' communication abilities.

Active Information Processing and Evaluation of GenAI Outputs

Selecting and prioritizing relevant technical information (OBS9) reached level 2 in 9/12 episodes (75.0%) and level 1 in 3/12 (25.0%). Students selected key terminology and content rather than incorporating entire AI outputs into their products. This behavior aligns with information selection in Mayer's (2009, 2021) CTML, but provides evidence of task practices rather than direct evidence of understanding or retention.

Checking AI-generated information (OBS10) reached level 2 in 10/11 episodes (90.9%) and level 1 in 1/11 (9.1%). Identifying and correcting inaccurate or unsuitable outputs (OBS11) reached level 2 in 6/10 episodes (60.0%) and level 1 in 4/10 (40.0%). These distributions suggest a distinction between cross-checking information and identifying and resolving inaccuracies. Checking outputs is necessary, but does not mean that learners have identified and correctly addressed every problem in terminology or visual representation.

Learner agency in GenAI use (OBS12) reached level 2 in 7/11 episodes (63.6%), level 1 in 3/11 (27.3%), and level 0 in 1/11 (9.1%). Recorded behaviors included adjusting terminology, images, or layout and rejecting unsuitable suggestions. Learners participated in decisions about product content and form, although agency was not equally evident across episodes.

Classroom observations showed that students used visual materials in presentations, selected relevant content, and checked GenAI outputs. Greater difficulties arose in choosing suitable visual signs, organizing information, and representing technical relationships. These findings point to the need for closer attention to how students explain and justify their visual choices.

4.2. Analysis of Student-Generated Infographics

The 12 infographics produced by two groups across six units combined English terminology with images, symbols, arrows, and spatial layout. The analysis focused on technical accuracy, coordination between text and images, and information organization. As the coursebook already included visual materials, students both transformed verbal content into images and reorganized existing diagrams.

Re-representing Technical Content

The products used different forms of representation suited to their content. The Exterior and Instrument Panel / Dashboard used labeled images to identify components; Active vs Passive Vehicle Safety used parallel columns to compare functions; and Four-Stroke Engine, Rack-and-Pinion Steering System, and Airbag System presented operating steps. These differences suggest that students had begun to select visual forms according to the knowledge being presented rather than applying the same layout to every topic.

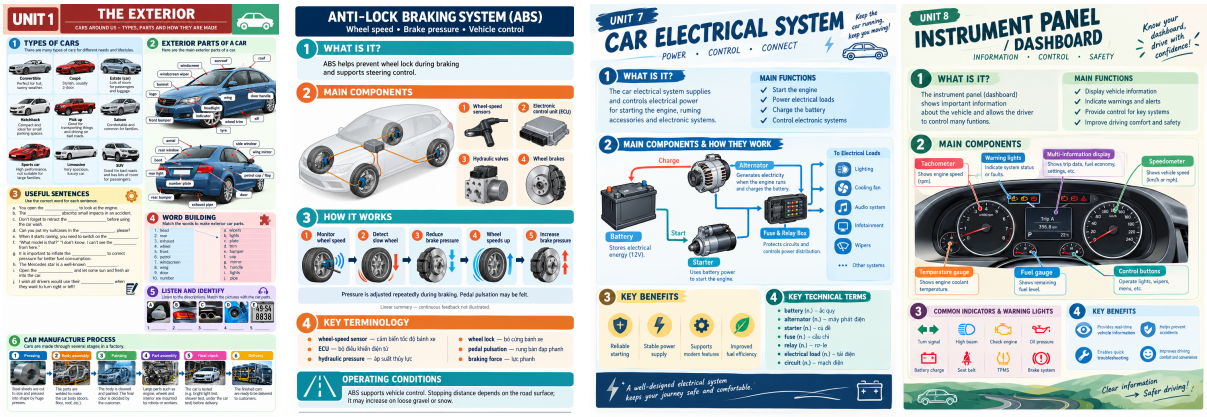


Figure 2. Student-generated infographics (1)

Technical relationships were not consistently preserved, however. In Rack-and-Pinion Steering System, the text described rotational and linear motion, but arrows around the wheel could be interpreted as rolling motion. *Car Differential* connected straight driving, cornering, and travel on uneven surfaces with successive arrows, potentially presenting distinct operating conditions as stages in a single process. Similarly, *Car Manufacture Process* presented "Body

Shop" and "Welding" as equivalent stages, without clearly showing the relationship between a production area and an activity performed within it.

In the simulated *Anti-lock Braking System (ABS)* infographic, the operating diagram presents pressure reduction and reapplication as a linear sequence without showing the continuous feedback between wheel-speed sensors, the ECU, and hydraulic control. Although the accompanying text mentions repeated adjustment, the visual structure does not convey this iterative relationship. This illustrates how a clear, concise layout can retain key terms while simplifying the technical relationships they represent.

Condensing content also reduced explanatory depth in some products. *Four-Stroke Engine* used "REPEAT 4 TIMES," which could confuse the four strokes with the number of cycle repetitions. *Car Electrical System* identified the functions of basic components but did not clarify features specific to electric or hybrid vehicle electrical systems. From an intersemiotic translation perspective, these examples show that content can be visualized without fully preserving its technical meaning structure.

Coordinating Modes in Meaning-Making

Terms were generally placed near images or connected to the relevant components, helping readers identify their referents. In *Rack-and-Pinion Steering System* and *Car Differential*, label colors matched connecting lines and descriptions; some labels also specified functions or types of motion. *Active vs Passive Vehicle Safety* combined colors, icons, and parallel columns to highlight functional differences. *Instrument Panel / Dashboard* placed warning symbols beside their English names, linking terminology to signs encountered in driving. Text and images thus complemented one another rather than merely appearing together on the page.

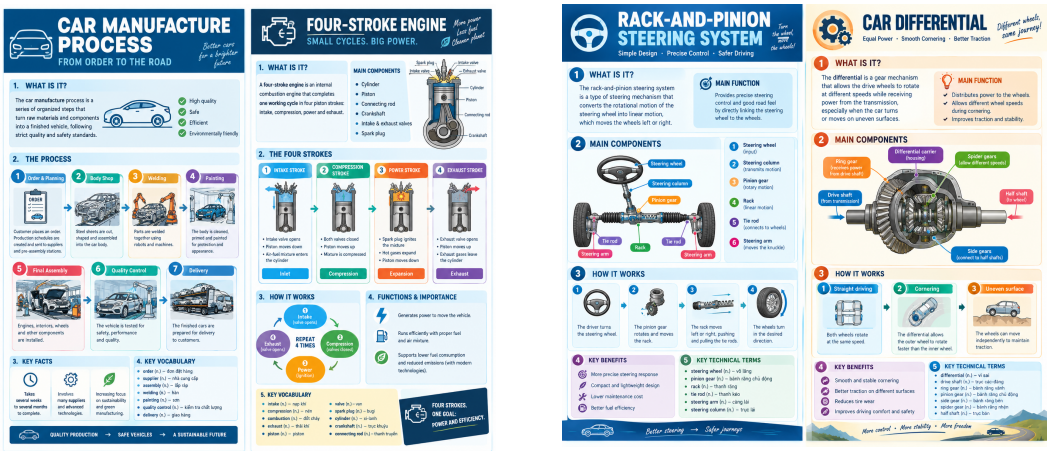


Figure 3. Student-generated infographics (2)

Some connections nevertheless remained inaccurate. In *The Exterior*, the line for "petrol cap / flap" did not point to the correct location; in *Four-Stroke Engine*, the intake arrow conflicted with the verbal description. *Airbag System* also contained inconsistent numbering: "Front airbags" was numbered 3 in the list but 1 in the illustration. From a social semiotic perspective, these mismatches matter because meaning is formed through relationships among labels and images as well as within individual elements.

Selecting and Organizing Information

The infographics commonly divided content into brief sections on definitions, components, operating principles, and terminology. This structure was evident in *Rack-and-*

Pinion Steering System, Car Differential, GDI Engine, Electric Power Steering and Airbag System, helping foreground key information and support oral explanations. Bilingual vocabulary lists in *Airbag System*, *GDI Engine and Electric Power Steering* and products on mechanical systems supported comparison of names, while function-labeled diagrams offered opportunities to connect terminology with technical knowledge. These features can be interpreted through CTML's processes of information selection and organization.

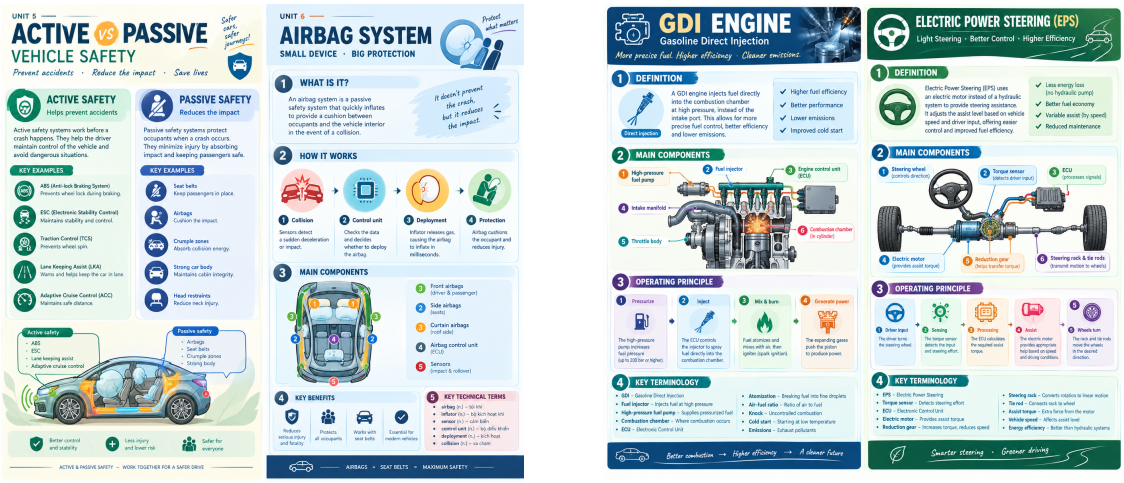


Figure 4. Student-generated infographics (3)

The infographics showed that students could link technical terms with images and organize content for presentations. However, inaccurate labels, misleading arrows, and unclear technical relationships remained in some products. This was consistent with the observations, where checking information was more evident than identifying and correcting errors.

4.3. Questionnaire Results on GenAI Experiences

Table 2. Students' Perceptions of GenAI Use in ESP Learning (N = 77)

Code	Item summary	M	SD	Agreement (%)
Q1	ChatGPT supports text condensation and terminology explanation	3.53	1.28	67.5
Q2	Canva AI supports transforming descriptions into diagrams and infographics	3.48	1.21	62.3
Q3	Combining both tools supports transforming concepts from text into images	3.65	1.29	72.7
Q4	Combining text and images supports fuller representation of knowledge	3.70	1.37	68.8
Q5	Visual elements support presentation confidence	3.57	1.30	59.7
Q6	Visual layout helps listeners follow the content	3.78	1.32	71.4
Q7	Combining text and images supports vocabulary retention	3.69	1.27	71.4
Q8	Content selection and diagram design support understanding	3.64	1.28	68.8
Q9	Visual representation reduces feelings of overload or stress	3.62	1.30	67.5

Note: M = mean; SD = sample standard deviation; agreement includes ratings of 4 and 5

The questionnaire obtained 77 responses to each five-point Likert item. Of the participants, 73 (94.8%) had used GenAI before the ESP course and four (5.2%) had not.

Mean scores for the nine items ranged from 3.48 to 3.78, with agreement and strong agreement ranging from 59.7% to 72.7%. Standard deviations of 1.21–1.37 indicate variation in responses despite an overall positive tendency.

Support for Processing and Transforming Technical Content

Students rated ChatGPT positively for summarizing technical content and explaining terminology (Q1: M = 3.53). Canva AI received a slightly lower rating for creating diagrams and infographics (Q2: M = 3.48), while combining the two tools attracted the highest agreement in this group (Q3: 72.7%). Students appeared to value the support these tools provided at different stages of the task, although the infographic analysis showed that some technical relationships were still represented inaccurately.

Multimodal Representation and Oral Presentations

Students generally agreed that combining text and images helped them represent technical knowledge (Q4: M = 3.70). Visual layout received the highest mean across the questionnaire (Q6: M = 3.78), while confidence in presenting attracted less agreement (Q5: 59.7%). This difference suggests that students valued visual materials for organizing their explanations, although they were less certain about their contribution to speaking confidence. These ratings reflect students' perceptions rather than direct evidence of audience comprehension.

Perceived Retention, Understanding, and Overload

The mean for Q7–Q9 was 3.65. Students positively evaluated support for vocabulary retention (Q7: M = 3.69; 71.4%) and understanding through content selection and diagram design (Q8: M = 3.64; 68.8%). These perceptions align with the role of selection and organization in Mayer's (2009, 2021) CTML, but remain self-reports rather than evidence of improved understanding or long-term retention.

Q9 received M = 3.62 and 67.5% agreement regarding reduced feelings of overload or stress. Because the item combined two different experiences, it cannot be treated as a measure of cognitive load. Similarly, referring to text and images as "two channels" in Q7 did not distinguish representational modes from cognitive processing channels; agreement does not validate the dual-channel mechanism.

Benefits and Difficulties of GenAI Use

The two open-ended questions yielded 57 responses about benefits and 56 about difficulties, written in Vietnamese. Reported benefits included saving time, explaining terminology, condensing ideas, and supporting design. The comment "I can transform the content and learn new words at the same time" suggests that some students linked product creation with language learning. One respondent, however, reported not using Canva AI, indicating uneven experience with the two tools. Quotations here were translated from Vietnamese.

Difficulties concerned inaccurate outputs, unclear prompts, limited operational skills, and gaps in technical knowledge. The response "It is difficult to distinguish what is right from what is wrong" suggests that awareness of possible inaccuracies does not necessarily imply the ability to evaluate them. This complements the observations: cross-checking was frequently

recorded, but the quality of judgment and revision still needed to be examined through the products.

Comments about free accounts reflected access limitations, while “I overuse it” and “I am dependent on it” indicated that some learners recognized problems in their own use. Three responses reported no difficulties or none so far; very brief responses lacking context were not interpreted further.

5. Conclusions and Recommendations

This classroom case study examined how automotive engineering students used ChatGPT and Canva AI to develop infographics for ESP learning. Students generally viewed the tools positively, while classroom observations and artifact analysis documented their involvement in selecting, organizing, and checking technical information. The infographics ranged from labeled component diagrams to representations of operating processes.

Students were generally able to connect terminology with images, but some products contained inaccurate labels, misleading arrows, or simplified technical relationships. These problems remained even when students were observed checking GenAI outputs. The findings suggest that verifying information and representing technical knowledge accurately are related but distinct aspects of the task.

Instructors should therefore guide students to check not only individual terms and images but also the relationships they represent, including component functions, operating sequences, and different operating conditions. Assessment should consider technical accuracy, coordination between text and images, and students' ability to explain their design choices. Comparing successive versions of infographics and questioning students about their revisions may help clarify how they evaluate GenAI-generated content.

The study was limited to one ESP course at one university. Its findings describe classroom practices, student products, and perceptions rather than measured improvements in vocabulary knowledge or retention.

Ethical considerations

Before the study began, the research procedures and measures to protect participants were reviewed and confirmed by the Department of Science, Technology and International Cooperation at the University of Transport Technology, Vietnam. The study was carried out as part of regular ESP classroom activities. All 77 students were informed about the purpose of the research and how their responses and learning materials would be used. They provided written informed consent before completing the questionnaire, including permission to use their responses and student-generated infographics for research and academic publication. Participation was voluntary, and students were assured that their decision to participate or decline would not affect their course grades. All research data were handled confidentially, and students' identities were not disclosed in the findings or published infographics. As the course teacher was also involved in the research, a second teacher participated in classroom observations and qualitative data coding.

References

Bibauw, S., François, T., & Desmet, P. (2019). Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of dialogue-based CALL. *Computer Assisted Language Learning, 32*(8), 827–877. https://doi.org/10.1080/09588221.2018.1535508

Hafner, C. A., & Ho, W. Y. J. (2026). Doing digital multimodal composing with generative artificial intelligence: Towards an integrated critical digital literacies model. *Language Teaching Research*. Advance online publication. https://doi.org/10.1177/13621688261459662

Ho, P. X. P. (2024). Using ChatGPT in English language learning: A study on I.T. students' attitudes, habits, and perceptions. *International Journal of TESOL & Education, 4*(1), 55–68. https://doi.org/10.54855/ijte.24414

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning, 38*(1), 237–257. https://doi.org/10.1111/jcal.12610

Jakobson, R. (1959). On linguistic aspects of translation. In R. A. Brower (Ed.), *On translation* (pp. 232–239). Harvard University Press. https://doi.org/10.4159/harvard.9780674731615.c18

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 537–550. https://doi.org/10.1177/00336882231162868

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

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

Mayer, R. E. (2021). Cognitive theory of multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 57–72). Cambridge University Press. https://doi.org/10.1017/9781108894333.008

Pham, M. T., & Cao, T. X. T. (2025). The practice of ChatGPT in English teaching and learning in Vietnam: A systematic review. *International Journal of TESOL & Education, 5*(1), 50–70. https://doi.org/10.54855/ijte.25513

Pham, V. P. H., & Le, A. Q. (2024). ChatGPT in language learning: Perspectives from Vietnamese students in Vietnam and the USA. *International Journal of Language Instruction, 3*(2), 59–72. https://doi.org/10.54855/ijli.24325

Vo, T. K. A., & Nguyen, H. (2024). Generative artificial intelligence and ChatGPT in language learning: EFL students' perceptions of technology acceptance. *Journal of University Teaching and Learning Practice, 21*(6), Article e811. https://doi.org/10.53761/fr1rkj58