

# GenAI in Learning, Teaching and Assessment

15

Instructors

Ms Yao Yao, Ms Tianyi Han & Prof. Dechao Li

Department

Language Science and Technology (LST)

FH6301

Theories of Translation

## Why did the instructor use GenAI for learning and teaching?

The core goal of the team led by Prof. Dechao was to prepare postgraduate translation students in Hong Kong for an AI mediated professional landscape, where machine translation (MT) and post editing (PE) are now standard practice for language service providers. Rather than treating generative AI (GenAI) as a threat or a 'black box' shortcut, the instructors designed a study to achieve the following:

- Help students develop MT literacy and post editing competence, both of which are highlighted as key professional skills in international translation frameworks.
- Expose students to a hybrid workflow in which GenAI supports, but does not replace, human judgement.
- Shift students from passive AI use (i.e., simply accepting MT output) to critical, informed engagement with AI-generated suggestions.
- Generate empirical evidence on how GenAI affects student effort—specifically time, editing behaviour, and cognitive load—under different task demands (full versus light post editing).

## How was GenAI used in this scenario?

The team implemented GenAI (GPT 4) within a carefully designed experimental PE task involving 26 MA translation students from three universities in Hong Kong. The key design features were as follows:

- **Task type:** Students post edited English–Chinese machine translations of short United Nations report excerpts (80–90 words), representing a realistic yet manageable genre.
- **Two post-editing modes:**
  1. Traditional Post Editing (TPE): Students were provided with the source text (English) and the machine translated output (Chinese). They revised the MT output relying solely on their own judgement and the source text.
  2. AI Assisted Post Editing (AIPE): Students were provided with the source text, the MT output, GPT 4's post editing suggestions, and GPT 4's chain of thought (CoT) reasoning (i.e., explanations of why particular changes were suggested, such as word choice, style, or accuracy issues).
- **Two translation briefs:**
  1. Full post-editing (FPE): Designed to achieve 'publishable quality' comparable to human translation; high intervention was expected.
  2. Light post editing (LPE): Designed to be 'good enough' for comprehension; minimal intervention was expected.

(Note: Each student performed tasks in both FPE and LPE, and in both TPE and AIPE modes, using a Latin square design to balance order effects.)
- **Data collection tools:** To understand the students' effort in a fine-grained manner, the instructors utilised the following:
  - Eye tracking: To measure where and how long students looked at the source and target texts.
  - Key logging: To record the number and type of edits, insertions, and deletions.
  - Pause analysis: To track how often and how long students paused while working.
  - Self rating and interviews: To capture students' perceived difficulty and post-task reflections.

## What was the impact on student learning?

The impact was nuanced rather than uniformly positive. Ultimately, GenAI changed how students worked more than how long they spent.

### 1. Time on task did not decrease

- Students did not work faster in the AI assisted mode.
- The average time per word remained similar across both TPE and AIPE.
- Many students reported that reading and evaluating GPT 4's explanations and suggestions was time consuming, especially when they ultimately chose to reject them.
- Pedagogical implication: GenAI does not automatically produce efficiency gains for learners; without targeted training, its 'help' can actually slow them down.

### 2. Technical effort increased with GenAI

- In the AIPE mode, students made more edits and insertions per word than in the TPE mode.
- GenAI suggestions appeared to embolden students to make finer stylistic and lexical adjustments and to engage more deeply with the text rather than settling for 'good enough'.
- Pedagogical implication: This increase can be viewed as productive effort. More keystrokes often indicate more thorough engagement and the potential for higher translation quality, even if it mechanically appears to be more work.

### 3. Cognitive effort showed mixed patterns

Objective measures pointed in two distinct directions:

- Less effort on source text comprehension: With GenAI support, students demonstrated shorter and fewer fixations on the source text. This suggests that GenAI helped them interpret the source text more quickly and confidently.
- More effort in decision making: Pause data revealed more frequent and relatively longer pauses in the AIPE condition. Furthermore, interviews indicated that students struggled with critically evaluating GenAI suggestions and deciding whether to trust the AI or their own judgement. Many students felt that using GenAI effectively required stronger bilingual proficiency and critical thinking, not less.

(Note: Subjective self ratings of effort did not differ significantly between AIPE and TPE, echoing known discrepancies between perceived and objectively measured cognitive load.)

### 4. Translation brief strongly shaped effort

Across both modes:

- Full post editing (FPE) consistently required more time, more edits, and greater cognitive effort than light post-editing (LPE).
- The benefits of GenAI regarding eye tracking measures (i.e., less time spent staring at both source and target texts) were more pronounced in light PE. For 'good enough' tasks, GenAI suggestions were easier to leverage. For high stakes, publishable tasks, students still needed to rely heavily on human judgement and additional editing.
- Pedagogical implication: GenAI was most helpful when quality expectations were moderate. For high quality briefs, it supported, but did not replace, intensive human intervention.

## What were the challenges encountered during the implementation and what solutions were used?

Overall, this case demonstrates that using GenAI in translator education are most powerful and less challenging when:

- It is framed as intelligence augmentation, rather than a human replacement.
- Tasks are closely aligned with explicit briefs and industry realistic expectations.
- Students are actively supported in developing critical AI literacy — the ability to question, adapt, and strategically utilise GenAI, rather than relying on it simply to speed up their work.

These lessons are highly transferable beyond the field of translation. In any academic discipline, GenAI works best educationally when it is embedded within a structured workflow, paired with explicit task parameters, and used to develop—rather than bypass—students' independent disciplinary judgement.