

## Navigating AI in Foreign Language Teaching: Balancing Innovation and Integrity in Course Design and Assessment

Kan Qian The Open University

25 April 2025
Online CI at OU & Lancaster CI Joint Teacher Training Day





## **Contents**

- Why should we teach students GenAl?
- What and how should students be taught about Generative AI?
- Critical Al Literacy Framework
- Examples from OU modules
- Findings from an OU GenAl assessment project
- Which assessment types are more robust
- Examples from OU's modules
- Clear and transparent guidelines and marking criteria
- What else can be done to minimise students' misuse of GenAl
- Next steps
- Some useful resources





## Why should we teach students GenAl?

"Generative AI has great potential to transform the Higher Education sector and provides exciting opportunities for growth. However, integrating it into teaching, learning, and assessment will require careful consideration." (A Department for Education spokesperson, BBC report 24 Oct 2024)

## Why should we teach students GenAI?

- Universities have a responsibility to teach students GenAI skills as it is a crucial future employability skill
- GenAl has great potential to support and enhance teaching & learning of languages and cultures

'If we don't educate [students] on how to use GenAI to help them in whatever material or whatever subject they're studying, they're going to be behind the rest of their cohort, their peers, when it comes to finding a job.' *OU tutor*  'Al skill has to be in the same way that we teach students to research. They've got to learn to keep up with Al...this is a new part of human existence if you want to be very fanciful' *OU tutor* 

## Risks: perpetuating stereotypes & bias

GenAI tools, like image generators, have been found to perpetuate harmful gender stereotypes. For instance, when asked to generate images of professionals such as doctors or CEOs, these tools often predominantly depict men, reinforcing the stereotype that these roles are male-dominated.

Language models can exhibit racial bias by generating text that reflects prejudiced views. For example, a study found that a language model developed by OpenAI, GPT-2, was biased against women, people of colour, and LGBTQ+ individuals. This can manifest in the form of negative or stereotypical descriptions of these groups.

Universities must determine how to harness the benefits and mitigate the risks to prepare students for the jobs of the future.

How to mitigate the risks?

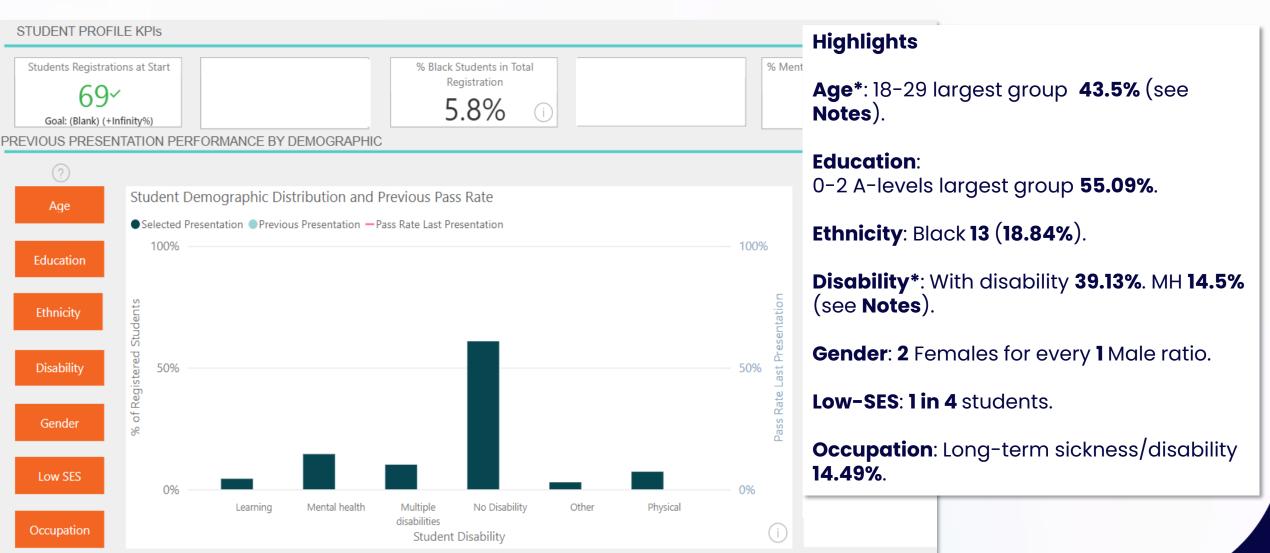
# What and how should students be taught about GenAl?

Adhering to OU's 2<sup>nd</sup> teaching and learning principle:

As far as is possible, the use of-Generative AI in Learning, Teaching and Assessment is anti-racist, anti-discriminatory, accessible and inclusive.

## L108 24J Student profile

Why EDIA is of particular importance at OU?



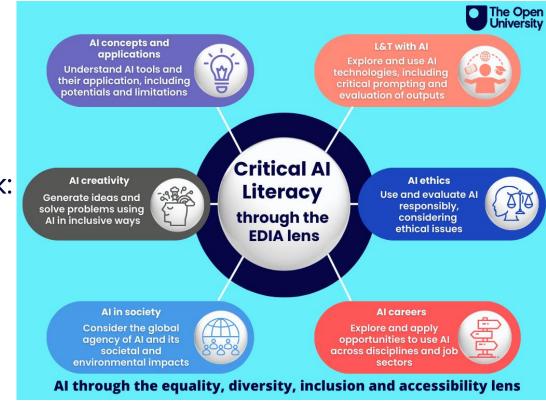
## What and how should students be taught about GenAI?

Students should be taught how to use GenAl responsibly and ethically, and our duty to raise their critical awareness with regard to EDIA concerns.

Critical Al Literacy Framework through EDIA lens
OU Al in LT&A Working Group, Prof Mirjam Hauck et al

The graphic illustrates the six areas of the framework:

- Al concepts and applications
- L&T with Al
- Al ethics
- Al creativity
- Al in society (especially environmental impacts)
- Al careers



#### Example of a language activity in Intermediate Chinese – evaluating outputs

林林的姑妈最近上了老年大学, ……她对她爸爸说:去上课就可以锻炼你的脑子。……林林告诉他大学有科技课,这门课就教怎么用APP。……

Below are three different versions of the English translation of the sample answer in Step B. One is Google Translation, one is a human translation, and the other one is Microsoft Copilot (it offers added data protection when logged in as an OU student). Use the blue pen to highlight the differences. Are the differences significant? Can you tell which version is the human translation? Put your comments below.

Version 1 (with line numbers)

Version 2 (with line numbers)

Version 3 (with line numbers)

#### Students then reveal the **Discussion** below with the table on the right:

Notice that only **Version 2** translated 就, the emphatic adverb meaning 'indeed' or 'actually'. It indicates that **Version 2** is most likely to be the human translation. Also Version 2's tense is inconsistent, which is another sign of human translation. The person who translated it also decided to put a quotation mark around 去上课就可以锻炼你的脑子 'going to attend classes may indeed exercise your brain', although it wasn't used in the original text.

Line No.	Chinese	English	English	English
(for		(Version 1)	(Version 2)	(Version 3)
English)				
2	爸爸	father	dad	father
2	让	encouraed	asked	asked
3	去上	attend	go to attend	go to
3	脑子不好	mind wasn't sharp	brain wasn't good	brain wasn't working
		anymore		well
4	去上课就	attending classes	'going to attend	going to class can
	可以锻炼	could help	classes may	exercise your brain
	你的脑子	exercise his mind	indeed exercise	
	10 10 10		your brain'	
6/7	这门课 <mark>就</mark>	that taught how to	which actually	this course taught how
	教怎么用	use apps	taught how to use	to use an APP
	APP		APPs	
7	报名	enroll	sign up	register
			<u> </u>	

## **Evaluating outputs (continued)**

Version 1 is translated by the Generative AI Copilot (accessed on 02/04/2025), and it reads very well with colloquial expressions such as 'mind wasn't sharp anymore'. One of the features of translations done by Generative AI tools is that they can produce translations that sound more natural and conversational. Another feature is that they can understand and incorporate broader context, making translations more nuanced and accurate. For example, when it translates 她对她爸爸说: 去上课就可以锻炼你的脑子, it turns the phrase after the colon ':' into an indirect speech by removing the colon and turn 你的脑子 into 'his mind'. However, if fails to pick up the meaning of 就, which is one of the tricky words in Chinese to translate.

Version 3 is by the Google Translate (accessed on 02/04/2025), and it is also near perfect but it tends to be more literal translation than colloquial in comparison with Version 1. Like Version 1, it couldn't' cope with the emphatic work 就. Also, when it translates 去上课就可以锻炼你的脑子, it uses the colon ':' in front of it straight from the original Chinese text and translates 你的脑子 into 'your brain', but does not use the quotation mark to signal it is a direct quote.

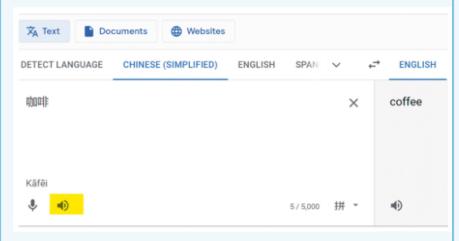
## Beginners' Chinese: using Google Translate for tone practice - its limitation

#### Step C

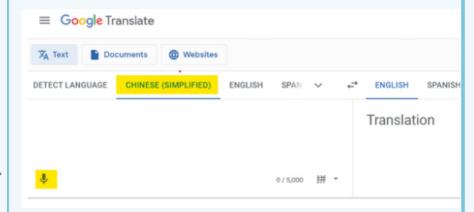
Practising tones is important because it will help you convey the meaning effectively when you speak Chinese. There is one technique that can help you learn the correct tone of words: the use of voice-recognition devices.

You are encouraged to try the following task, where you will pronounce individual Chinese words to an online translator, to check whether the voice recognition 'understands' the words you say. But you need to be aware that voice recognition devices may not be perfect and can sometimes fail to recognise what native speakers say too! So, it's important to see this as a repetition task that enables you to listen to a word many times, and repeat it many times, regardless of whether the program recognises it or not. Don't worry if, after trying several times, you don't get 'understood' by the machine; at least you will get used to hearing the tone of these words and you will be bound to improve your pronunciation. A Chinese speaker, in the right context, would be much more likely to understand you.

Go to an online translator that allows voice recognition, such as Google Translate. Set the language to be translated as Chinese. Copy and paste the words or phrases below (one each time). Listen to the word first, by clicking on the 'speaker' icon (see image). Pay special attention to the tones (you can see the tone marks on Pinyin transcription), but also to the sounds of consonants and vowels.



Once you have heard the word a few times, click on the microphone icon (see next image) to record yourself. Try and imitate the sounds and tones that you heard before



Screen capture of an online translator site highlighting the microphone icon

Show description

- 1. 咖啡
- 2. 夹克
- 3. 吉他
- 4. 你好
- 5. 再见
- 6. 谢谢

Did the voice recognition device (eventually) understand you? It is not unusual for the device to come back with a completely different word, especially at first, when the pronunciation of some sounds and tones seems more difficult to master. But rememb that it is not necessarily your fault if the system did not recognise your words.

What was the funniest translation that the machine returned when you tried to pronounce these words and phrases? Share your experiences in the space below.



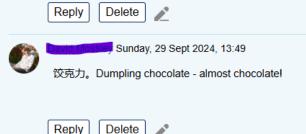


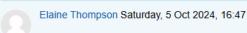
11 of 11

Collapse all posts

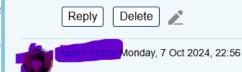
2. 吉他 car wash and locomotive instead of guitar

3





清蒸鱼 Qīngzhēng yú - steamed fish - translated as Fish steamed in a hot bath.



1. 夹克 - 贾克 Jacques rather than Jacket

2. 吉他 - 其他 Other rather than Guitar

Screen capture of an online translator site highlighting the speaker icon

## Using phone app to translate images into text –evaluate outputs

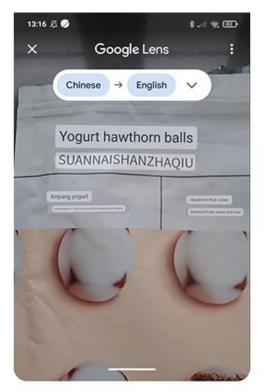


#### Translating texts from images with a phone app

A mobile phone application of particular interest for students learning Chinese (or any other language) is one that translates text that is in a photo taken with the phone's camera. So, for example, you can take a photo of a street sign, a poster or a leaflet, a memorial or gravestone, or even a packet of sweets, and see the image with the original text – in this case in different characters – substituted by a translation.

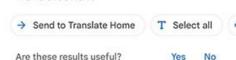


The app will automatically convert that picture into a cloned one, with all Chinese text translated to English (see below). You can also choose other target languages.



Translated text

Show description >



Make students aware often literal translation is provided by Google Translate – 'balls' in this context means 'sweets'.

## Teaching prompting, reflective and critical analysis

The following prompt was submitted to Gemini on 8/1/2025:

You are a mediumly successful undergraduate student. Please turn these bullet points into a short coherent paragraph.

- \* environmental situation = challenges and opportunities for Illy coffee
- \* climate change = hotter temperatures and changed rainfall
- \* impact on bean quality and yield?
- \* impact on supply chain and farmer livelihood?
- \* how to solve?

It provided the following response:

Climate change, with its rising temperatures and altered rainfall patterns, presents both challenges and opportunities for Illy Coffee. Increased heat stress can negatively impact bean quality and reduce yields, while unpredictable rainfall can lead to increased susceptibility to diseases and pests. This disruption in production significantly impacts the supply chain and threatens farmer livelihoods. To address these challenges, Illy can explore strategies such as supporting agroforestry practices, investing in climate-resilient coffee varieties, and promoting sustainable farming techniques that enhance carbon sequestration and biodiversity. These initiatives can not only mitigate the negative impacts of climate change but also position Illy as a leader in sustainable and ethical coffee production.

Question: use the text box below to note any cause and effect language you spot in the text.



[answer: can negatively impact and reduce; can lead to; impacts and threatens; to address these challenges; can mitigate and position]

The output of the GenAl provides some basic information, which could work well for a general audience such as that of a newspaper or magazine article. A more specialised business audience, however, would require a more informative and precise text. Use the text box below to note any improvements you think might be made to this text to become more suitable for a specialised audience.

[answer: whatever the student thinks]

As noted throughout this module, a text can be written differently to suit the audience, purpose and context, though the information may be the same. Comparing two paragraphs written differently can help you to notice differences.

## Teaching prompting and reflective and critical analysis (continued)

Gemini AI was asked to edit the paragraph; the first part of the revised answer is given below:

Climate change, with its rising temperatures and altered rainfall patterns, presents both challenges and opportunities for Illy Coffee, particularly in key coffee-producing regions like Central America, South America, and Africa. For example, a study by the Intergovernmental Panel on Climate Change (IPCC) found that rising temperatures can lead to increased heat stress on coffee plants, impacting bean flavor and potentially reducing yields, especially in regions like Colombia and Ethiopia, where coffee production is highly dependent on specific temperature ranges. Additionally, unpredictable rainfall can create conditions favorable for the spread of diseases like coffee leaf rust, further impacting production, particularly in regions like Central America where this disease has had significant impacts on coffee yields. This disruption in production significantly impacts the supply chain and threatens farmer livelihoods. To address these challenges, Illy can explore strategies such as supporting agroforestry practices, which can provide shade and improve soil moisture, in regions like Mexico and Guatemala, where agroforestry systems are already being implemented.

Question: Compare the original and the revised paragraphs. Post on the in-page forum noting what changes were made to the text and whether they are the same kinds of things you would have added.

[answers could include adding reference to a specific study, geographical locations, specific diseases]

[in the discussion for the activity, note what this means for AI, and whether there is enough detail etc]

## Learning design through an ethical lens

#### Responsible by design

OU's Learning Design team has developed a framework that looks at AI through an ethical lens and helps module teams to create and review their content.

The framework is centred around four pillars with reflective questions to review module content.

It focuses on ensuring AI is:

Fair Transparent Accountable Safe



#### **Environmental impact**

Consider the carbon emissions, computing power and energy consumption associated with GenAl.



#### **Encoded Bias**

Consider the normative predictions of AI models and how they mirror, reproduce and even exacerbate stereotypes.



#### **Exploitation**

Consider the exploitation of labour in the global south, as well as the use of huge datasets for which rights have not been obtained.



#### **Digital Divide**

Consider who has access to these tools, what level of access they have, and what implications this has for equitable technology use.

#### What is it?

The Responsible by Design (RBD) Framework is a tool that can be applied to any use of, or introduction to, GenAl in OU modules. It has been designed to prompt consideration of ethical implications.

#### Why is it important?

By considering GenAl through the lens of ethics, the RBD Framework will support wider OU strategy and facilitate a greater consideration of how we approach GenAl ethically and responsibly with our students.

## Teach how to use the tool responsibly

Design a nutritious food packaging using some of the 4-character expressions provided below. Choose the food items you have learnt so far.



Share your final finished design in the Activity forum as an attachment <link to the forum>. It can be a photo of your own drawing. If you use any GenAI tools, describe the tool used, the prompts used and the date it is accessed. Do bear in mind that every time you generate an image using a powerful AI model, it takes up as much energy as fully charging your smartphone.

# Findings from an OU AI & Assessment project



## Developing robust assessment in the light of Generative Al

OU's large-scale study funded by NCFE on the robustness of assessment in the light of Generative AI.

#### **Objectives:**

- •To identify the most robust and easiest assessment types to be answered by GAI, to enable institutions to risk-assess their assessments and identify the highest risk assessment(s) which may require changes.
- •To enable some comparison and sharing of best practice across subject disciplines and levels.
- To assess the effectiveness of a short training programme to upskill educators in recognising scripts containing Al-generated material.

#### **Methodology**

45 markers were recruited to mark 944 scripts, a mix of original student scripts and scripts generated by AI. The reasons given for GenAI detection were key data included in the qualitative research.

## What are the more robust types of assessment?

The AI answers for 58 of the 59 questions received passing marks (exception: an activity plan question which required specific application of module material supported by unambiguous marking guidance). GAI scored particularly high in Level 3

The 17 assessment types (assessed over 59 questions) used in this research were generally not robust in the face of GAI; either the GAI answers performed well and achieved a passing mark, or marker training increased the number of false positives.

## The following types are more robust.

- audience-tailored
- observation by learner
- > reflection on learning and work practice (all are what known as 'authentic assessment')

## Examples: assessment tasks in languages

**Role-Playing Scenarios**: Students can participate in role-playing activities where they simulate real-life situations, such as ordering food at a restaurant, asking for directions, or shopping at a store. This type of assessment allows students to use the target language in a practical context, demonstrating their ability to communicate effectively in everyday situations.

**Personalised Projects**: Students can create projects that reflect their interests and experiences, such as a photo diary, a simple blog, or a short video presentation about their daily routine or personal steps taken to reduce carbon footprint. These projects encourage students to use the language creatively and meaningfully, showcasing their understanding and ability to express themselves in the target language.

**Evaluating different translation outputs:** focusing on certain language structures and use.

Consider multi-stage assignment (e.g., outline assessed in one assignment → introduction assessed in another assignment → final piece with previous included but taking into consideration of teacher feedback)

## Interactive assessment: designed for learning and reflection

#### TMA01 consists of two parts:

- Part 1: an informal 15-minute online catch-up with your tutor in an online room, during
  which you will say a few sentences in Chinese and have a discussion in English about
  your experience of learning Chinese so far.
- Part 2: the submission of a written pro forma in English after the meeting.

The assignment is designed to be supportive. Marks in this assessment are not awarded based on how well you perform in your spoken Chinese nor the quality of the written proforma. You will be awarded **full marks** for

- 1. attending the meeting and
- 2. submitting the completed pro forma after the meeting.

#### It is therefore an opportunity

- · to practise and improve your learning
- to meet and work with your tutor on a one-to-one basis, so they can offer personalised and focused feedback and support
- · to focus on reflection about your learning

rather than a formal evaluation of what you have learned so far.

## Real Time Student Feedback Survey of the pilot

25 students took part in the survey (out of 55 who completed the assignment, 45%)

How could your first TMA experience be improved?

→ 20 responses: **12** of them overtly expressed enjoyment and satisfaction (even though that was not part of the question):

a really good experience / excellent and I thoroughly enjoyed it / [the tutor] made me feel at ease and comfortable / I found it very useful / I really enjoyed the TMA01 session!

#### **SPECIFIC POSITIVE OUTTAKES:**

- it is beneficial to have the speaking TMA at the beginning of the module [...]
- It was nice to get some one on one feedback without feeling as if it was a test this early on.
- The reflection before the TMA is a good exercise
- The 1:1 time with a tutor was appreciated given the focus on self-study which comes with OU degree.

## **Examples in Education**

#### Personalised educational case studies

Implementation: Assign each student a unique case study focusing on different educational settings, policies, or challenges, requiring analysis and solution development.

Example: A Level 3 student analyses an inclusive education policy for a rural school, proposing strategies to improve access and engagement.

### Al-assisted but critical evaluation of curriculum design

Implementation: Allow students to generate AI-assisted lesson plans or curriculum frameworks but ask them to critique and adapt them based on pedagogical theories.

Example: A Level 3 student evaluates an AI-generated literacy curriculum for primary schools, refining it to support students with diverse learning needs.

## Clear and transparent guidelines

#### **OU's guideline for students re GenAl**



#### Generative AI and assessment

While the use of Generative AI for learning is unrestricted, its use for each formative or summative assessment piece (e.g. a single question, or an MSc or PhD dissertation) falls into one of three categories.

#### Category 1: You cannot use Generative AI to complete an assessment piece.

Some assessment questions are designed to help you practice and demonstrate basic knowledge and skills (including the application of concepts, developing arguments and following procedures) needed for your study and employment. This category typically applies when the use of AI tools would prevent you from acquiring those skills. Using Generative AI for this category is subject to academic conduct procedures.

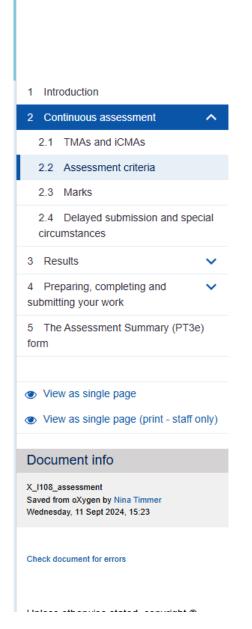
#### Acknowledging Generative Al

If the assessment falls into Category 2 or 3, you must acknowledge the use of Generative AI, except for minor changes to the form of your answer. This means that you don't need to indicate the use of AI tools for correcting spelling and grammatical mistakes or for converting passive to active voice, for example. However, any Generative AI used for producing the content of your answer, or for substantial changes to the form, must be acknowledged. Examples include ideas and outlines, rephrasing and restructuring paragraphs, and changing informal to academic style.

Unless the assessment tells you otherwise, you should report the use of Generative AI as follows.

- Reference all Al-generated outputs, using the Generative AI (Harvard) style. Use of third-party material (including AI outputs) in assessment without citing it constitutes plagiarism, which is an academic conduct offence.
- Summarise how you used Generative AI in an appendix that describes your conversation with the tool and what you did with its outputs. Put your prompts in quotation marks and cite the conversation as personal communication. For example,

## Every module has an Assessment guide with clear criteria



#### 2.2 Assessment criteria

As TMA 01 is designed for practice and personalised feedback from the tutor, and for encouraging you to reflect on the learning of Chinese up to that point, you get the full mark so long as you satisfy the following conditions i) attending the catchup meeting with your tutor; ii) completing 'TMA 01 submission pro forma'; and iii) submitting the completed pro forma via eTMA system by the cut-off date. This means that the quality of your language competence and the reflection in the pro forma are not assessed. For details of TMA 01 briefing, go to TMA 01 in the Assessment page.

TMA 02 and TMA 03 are marked according to how well the following criteria are met:

#### Assessment criteria for the writing task of TMA 02 and TMA 03

#### 1 Written communication and task completion

All aspects of the task have been completed in the format required and in accordance with the task brief and have been communicated to a good standard for this level.

#### 2 Accuracy of language and character use

Language structures appropriate to the level and task are used accurately throughout. Characters chosen for the task are accurate throughout.

#### 3 Range of appropriate language

A range of language structures, vocabulary and expressions appropriate to the level and task are used throughout.

#### Assessment criteria for the speaking task of TMA 02 and TMA 03

#### 4 Spoken communication and task completion

All aspects of the task have been completed in the format required and in accordance with the task brief, and have been communicated to a good standard for this level.

#### 5 Range and accuracy of spoken language

A range of language structures, vocabulary and expressions appropriate to the level and task are used accurately throughout.

#### 6 Pronunciation, tones and delivery

## Student engagement with the Assessment guide



## What else can be done to minimise student misuse of GenAl?

- > Identify the motivation of misuse: stress caused by workload, fear of failure, deadline
- How to mitigate: reassess academic workload and reduce it if too heavy; offer study support session, resit and extension; design assessments that minimise last-minute panic, such as scaffolded/multi-layered submissions (e.g., requiring outlines first, then drafts and the final submission, incorporating teacher feedback from the previous stages).
- > Understand possible justifications? : others are using it so I may just as well; I'm only doing it to improve the quality of my assignment; it is not really cheating as I do some post-editing...

How to mitigate: clear AI use guidelines in assessment; teach academic integrity; integrate reflective engagement in assignment and assessment (as mentioned earlier) which requires students to briefly explain how they used AI in assignments, assessing its benefits and limitations.

> Understand the dilemma of escaping the punishment and being punished?: As some assignment tasks can pass by easily without leaving much track of GenAI, students are more tempted; but at the same time some students are worried about being punished by hoping that they won't get caught.

How to mitigate: redesign assignment task to avoid summary-based or direct translation tasks; re frame Al use as an acknowledged tool: i.e. in the assignment briefing, make it clear that they must disclose Al assistance and justify its use, and critically evaluate it; also give marks to source justification.

## **Moving forward**

- Universities also need to focus on teaching the human and soft skills which GenAl cannot replicate (e.g. reflection and critical thinking, academic integrity)
- Sharing good practices and examples of AI-assisted language learning activities and exam questions
- Carry out research in this field to better understand how students use GenAl and machine translation tools
- Institutionalised continuous training programmes for the teaching staff
- Regularly review guidelines regarding GenAi and assessment and be transparent
- Develop online short courses and resources

## Al related online short courses at OU

#### An Introduction to Artificial Intelligence (LG003, 40 hours of study)

This short course critically examines the historical, ethical and social issues around this technology within society (https://www.open.ac.uk/courses/short-courses/lg003)

#### Al language technology in the workplace (LG007, 20 hours of study)

This short course explores the rapidly rising forms of Artificial Intelligence (AI) such as language technologies that use human language to carry out daily activities, communicate with people and control systems and devices. You'll examine the impact on our everyday working lives, as well as the broader historical representation of society through this technology.

#### Teaching Languages and Cultures with AI (forthcoming in May 2025)

This forthcoming short course focuses on how to utilise AI effectively for the learning and teaching of languages and cultures with a focus on equitable and inclusive approaches. To mark the lunch of this short course, we have a public lecture 'Unlocking creativity: the importance of

Al Literacy' on 27th May (1-2pm). Sign up via Eventbrite.

Mode: Enrol anytime. Self-paced, online study with learning support adviser, digital badge for attendance



## OpenLearn – OU's free-learning platform (open.edu/openlearn/)

➤ All my own work: exploring academic integrity (short course which contains one block on GenAl) https://www.open.edu/openlearn/mod/oucontent/view.php?id=148128&section=4.1

Integrity
The course has now been added by your OperiLeam Proble. There you can track your progress and find your statement of participation.

SESSION

SESSION

SESSION

3.1 Generative artificial intelligence
In the previous section, you expired some of the templations Grace faced when preparing her assignment. There are others, of course, and in this section you will consider one more closely: generative artificial intelligence (client).

➤ Machine translation – how teachers can deal with online translation tools in language studies assessments

https://www.open.edu/openlearn/education-development/machine-translation---how-teachers-can-deal-online-

translation-tools-language-studies-assessments

> Creativity and artificial intelligence

https://www.open.edu/openlearn/languages/linguistics/creativity-and-artificial-intelligence





## 13-14th Nov 2025 EuroCALL GenAl Conference



HOME NEWS ABOUTUS - SIGS - EVENTS - PUBLICATIONS - JOIN

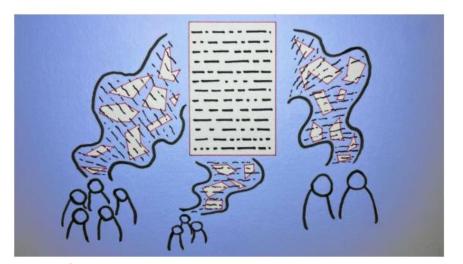
#### GenAI Conference @ The Open University UK

Home » EVENTS » GenAl Conference @ The Open University UK

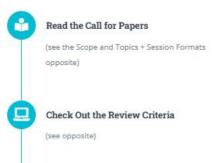
#### Shifting Power in Language Learning and Applied Linguistics with GenAI

A conference at The Open University UK, School of Languages and Applied Linquistics, in Milton Keynes and online

November 13th and 14th 2025







# Thank you 谢谢!

Contact: <u>online-ci@open.ac.uk</u> <u>ci@lancaster.ac.uk</u>









