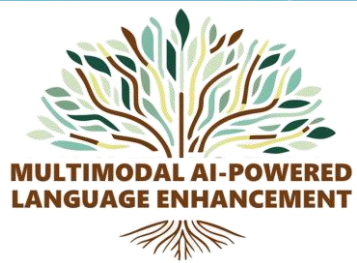


Learning & Teaching Expo 2024  
學與教博覽2024

11 - 13 Dec 2024

Hall 3B-E, Hong Kong Convention and Exhibition Centre

Shaping Education for  
a World of Change



# Multimodal AI-Powered Language Enhancement System 多模態人工智能驅動的語文能力增強系統



EdUHK 30th Anniversary  
教大三十周年

香港教育大學  
The Education University  
of Hong Kong

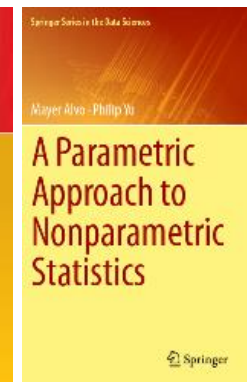
Professor Philip L.H. Yu

Department of Mathematics and Information Technology  
The Education University of Hong Kong

12 Dec 2024

# Philip L.H. Yu 楊良河

- Professor, Department of Mathematics and Information Technology, EdUHK
- Associate Director, University Research Facility of Data Science and Artificial Intelligence (數據科學與人工智能中心實驗室), EdUHK
- 2 research monographs and  $\geq 140$  research publications
- **Research Specialties:**
  - Preference Learning; Analysis of Ranking Data; Nonparametric Statistics.
  - Data Mining, Multimodal AI and its Applications to Education & Healthcare.
  - Time Series Analysis, Financial Data Analysis and Risk Management.
- **30+ Years of Multidisciplinary Research Experience:**
  - Working with experts in Language (EdUHK), Healthcare (HKU Med, Hospital Authority, DH), Finance (HKU, HKUST, HKBU, HKAB, HKEx, HKMA, BoCG, Manulife (HK)), Environment (CityU), Public bodies (FSD, MTR, EPD), etc.





# Background: language learning challenges

---



**Traditional classroom learning**  
(heavily relies on teachers,  
textbooks, and schools)



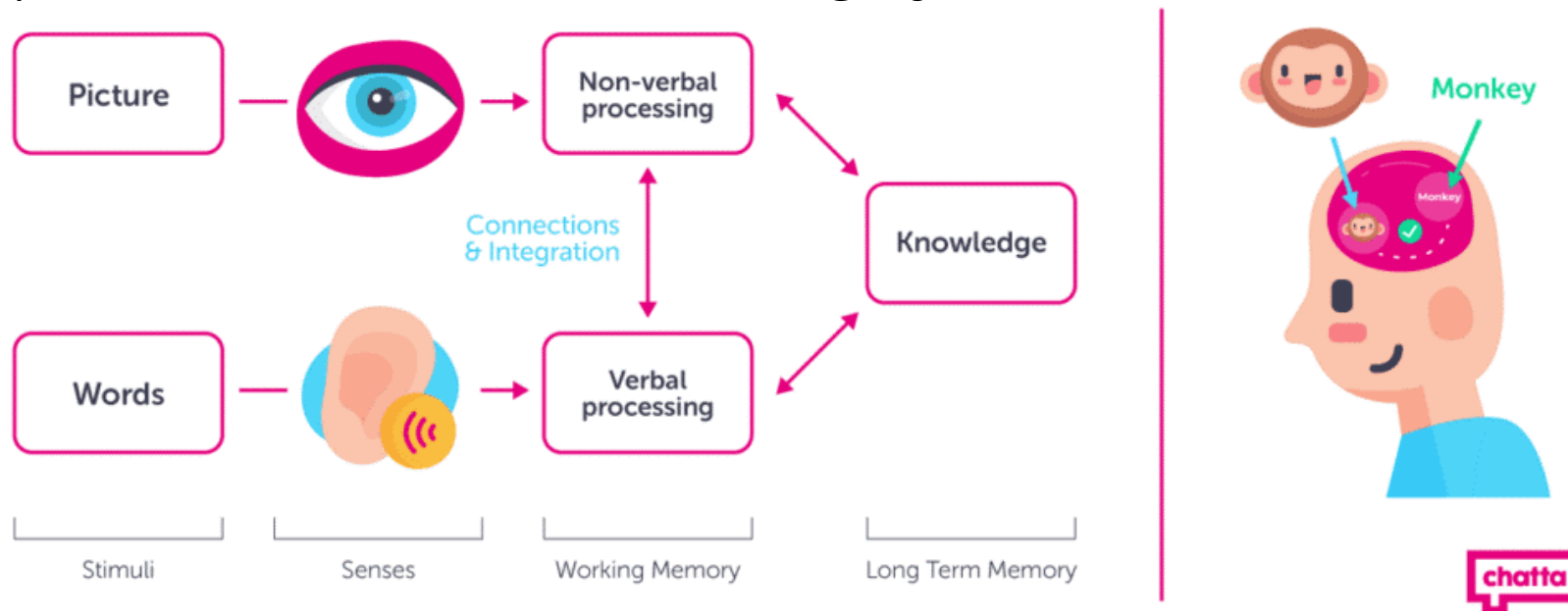
**Limited learning environments**  
(such as the communities with non-  
English speaking parents)



**Boring learning experience**  
(especially those students who  
lack interest in English learning)

# Dual-Coded Multimedia Language Learning

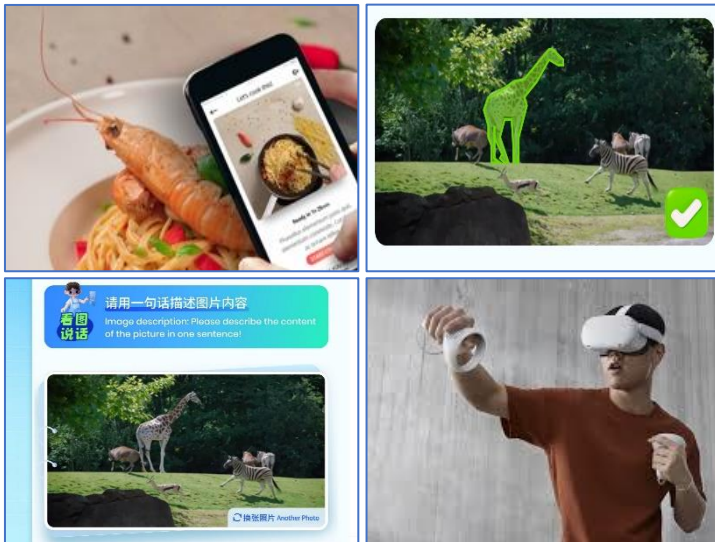
According to the theory of **Multimedia Learning**, information is better memorized when presented **both verbally and visually** because learners can receive information through both verbal and nonverbal sensors and utilize corresponding working memories to process the target information (Mayer, 2001). Wang and Li (2019) successfully developed a dual-coded multimedia application for improving reading comprehension of students with imagery deficit.



# Our Goals and Vision for Language Learning

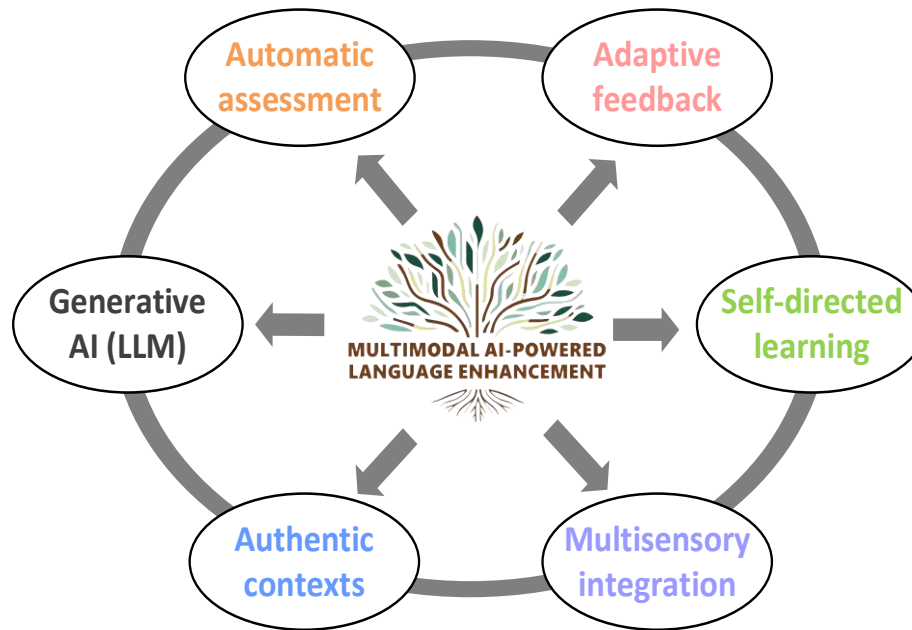
Our goal and vision are to utilize AI technology to empower students in self-directed learning English in authentic contexts through multisensory perception and interaction.

## Authentic situations & daily-living activities



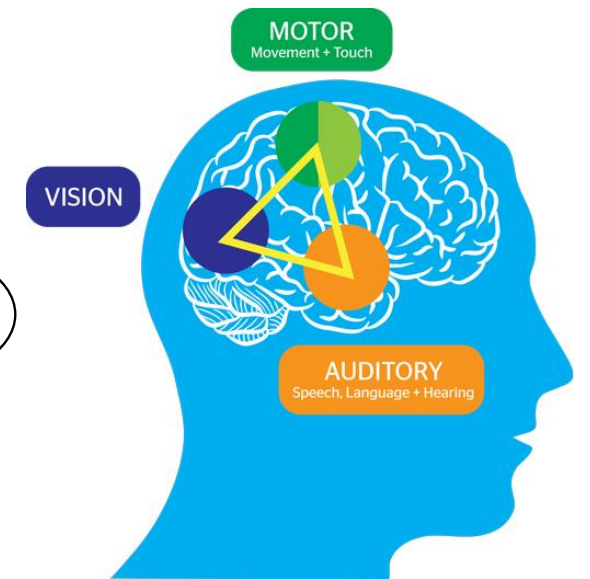
## Ubiquitous learning

## Automatic assessment & adaptive feedback



## Self-directed learning

## Multisensory perception & multimodal interaction



## Effective learning

# Picture-cued Writing Tasks

---

Picture-cued writing tasks are capable of presenting information in an intuitive and objective way, which are suitable for young writers with low levels of literacy skills.

**A typical picture-cued task asks students to describe a given picture**



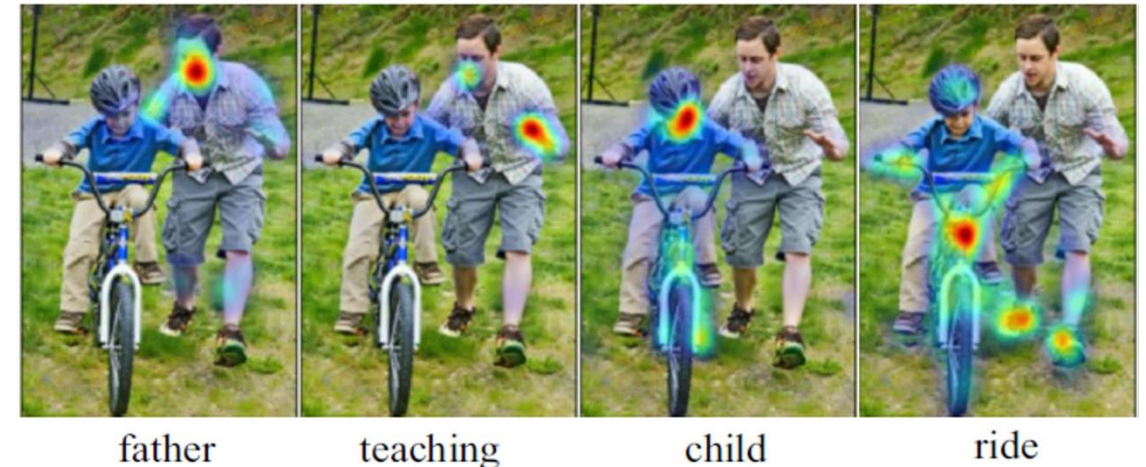
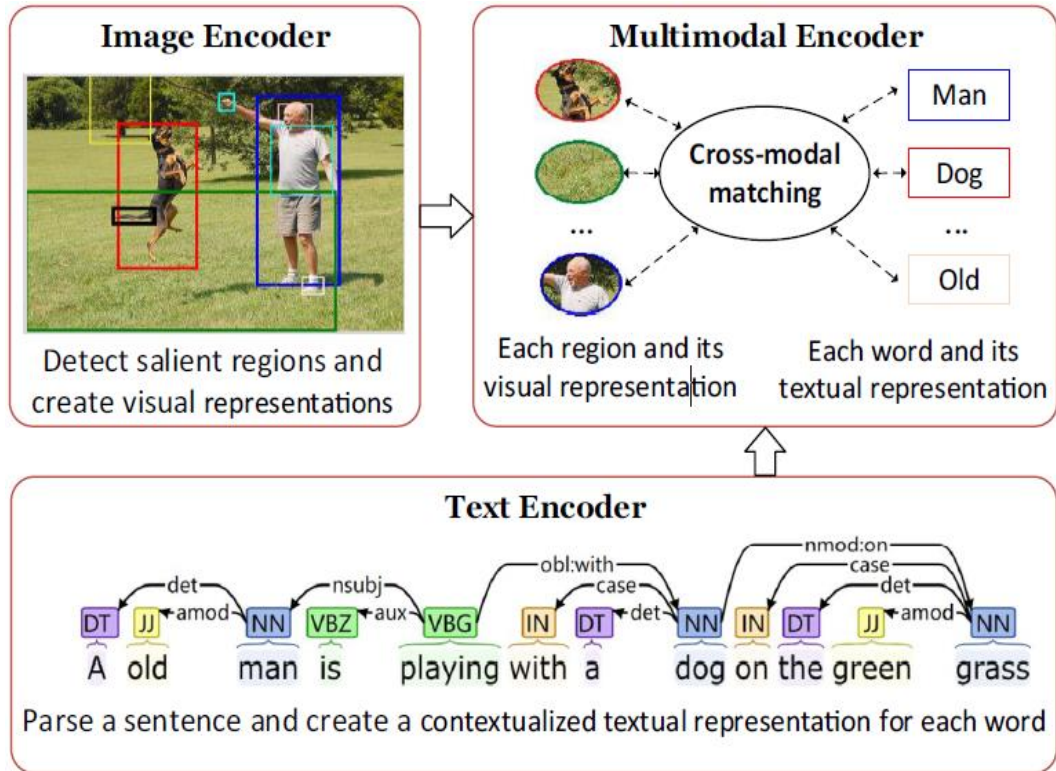
Student A: “The boy is learning how to ride a bike.”

Student B: “On the grass, boy is trying to keep his balance on the bike, and his father is following him to ensure his safe.”

**On a scale of 0 to 10, what scores do they receive?**

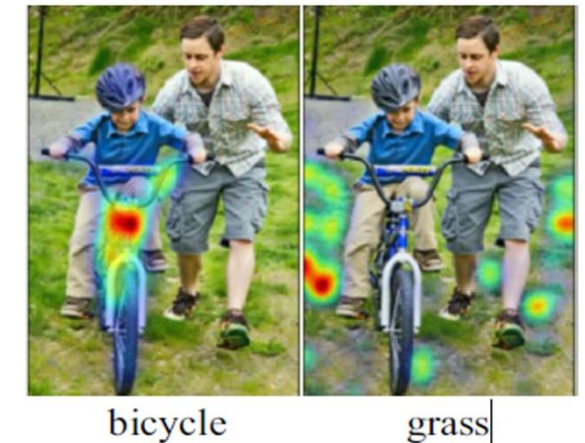


# Predictive Models for Automated Scoring



## Six dimensions:

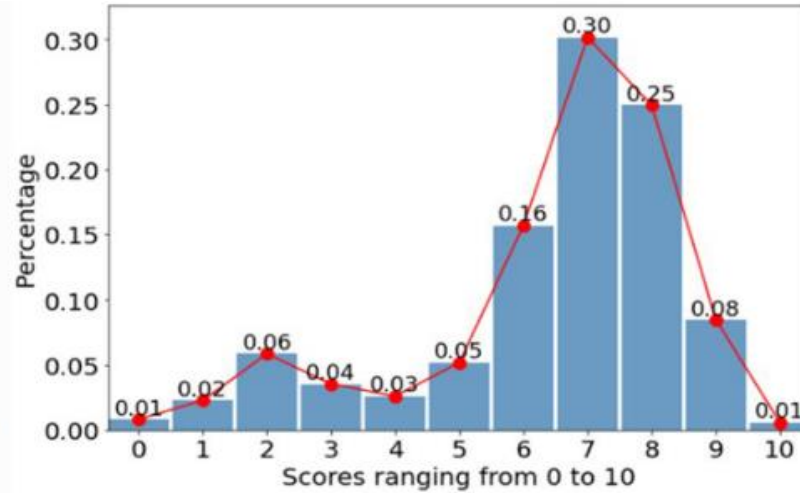
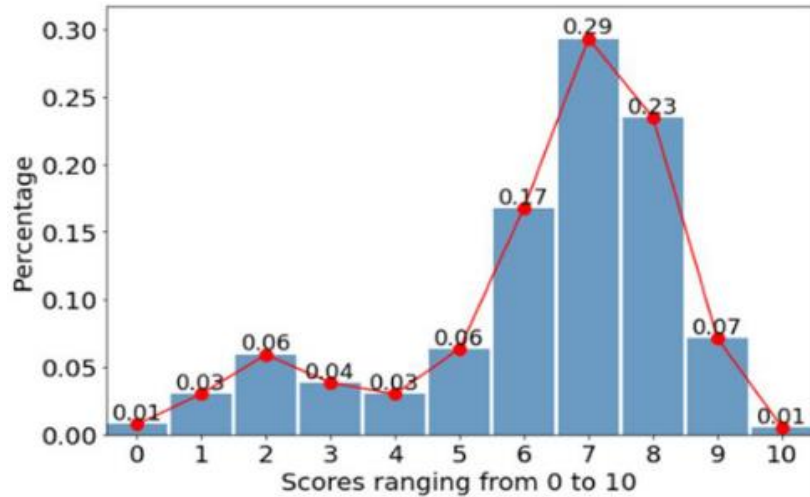
- Grammar [0, 3]
- Spelling [0, 1]
- Convention [0, 1]
- Comprehensiveness [0,3]
- Vividness [0, 1]
- Sentence structure [0, 1]



Cross-modal deep learning for automated scoring

Zhao, R., Zhuang, Y., Zou, D., Xie, Q. and Yu, P.L.H. (2023). AI-assisted automated scoring of picture-cued writing tasks for language assessment. *Educ Inf Technol.* 28(6), 7031–7063.

# Predictive Models for Automated Scoring

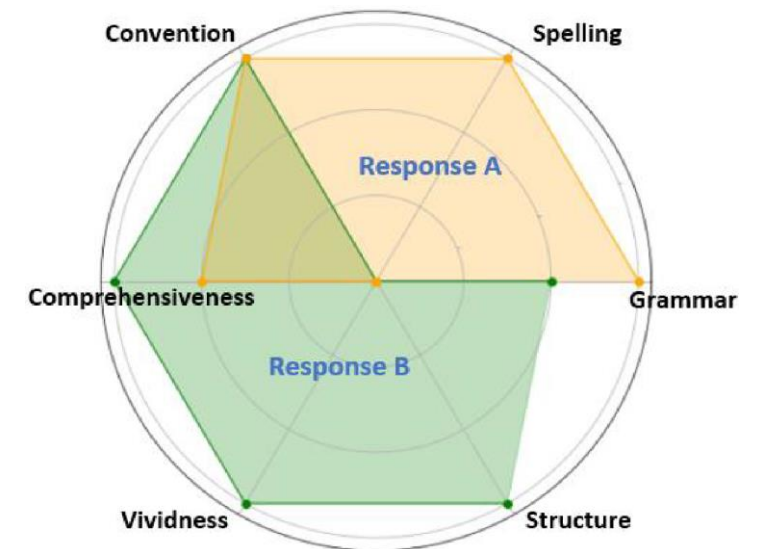


human scoring and our AI automated scoring

Two example students responses:

Response A: “The boy is learning how to ride a bike.”

Response B: “On the grass, boy is trying to keep his balance on the bike, and his father is following him to ensure his safe.”





# Automated Scoring and Automated Feedback



香港教育大學  
The Education University  
of Hong Kong



Department of  
Mathematics and  
Information Technology

 zhaotest

Questionnaire

Logout

Clear for rewrite

Submit for AI automated scoring

The website for supporting picture-cued writing

## AI自动产生的评分、评语和改进建议

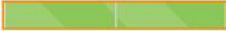


您写的句子: A man is playing with his dog on the grass.

AI评分 7.4 分, 共 10 分: ★★★★★★☆☆

**AI智能** 您的描写很棒! 如果更全面、准确、详细地描写图片中的人(物)、动作、场景及细节特征等, 您将获得更好的分数, 可以再挑战一次!

### 各项评分结果及修改建议, 根据AI建议, 修改句子、挑战高分!

- 图片内容[2.0 / 3.0 分]:  请仔细看看, 图中还有什么? 在哪儿? 干什么?
- 细节特征[1.0 / 2.0 分]:  多用准确的形容词、副词、介词, 得分会更高哦!
- 语法拼写[2.0 / 2.0 分]:  语法、单词拼写很不错!
- 单词使用[1.5 / 2.0 分]:  想一想, 有没更准确的名词、动词!
- 句式结构[0.9 / 1.0 分]:  不错, 句子结构完整, 试试更复杂点的句子!

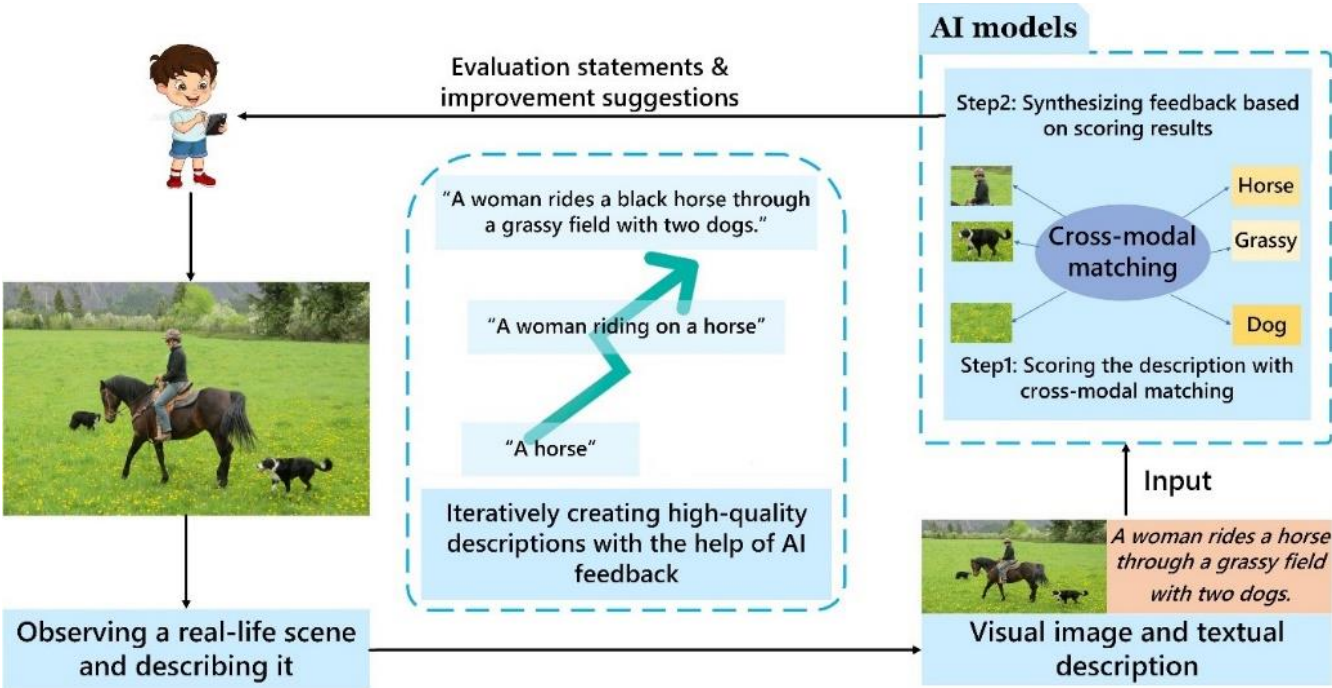
不再重写, 进入下题

再写一次, 挑战高分

A demo to illustrate the automated feedback provided by the website

# Feedback Help Students Progressively Improve Their Writing

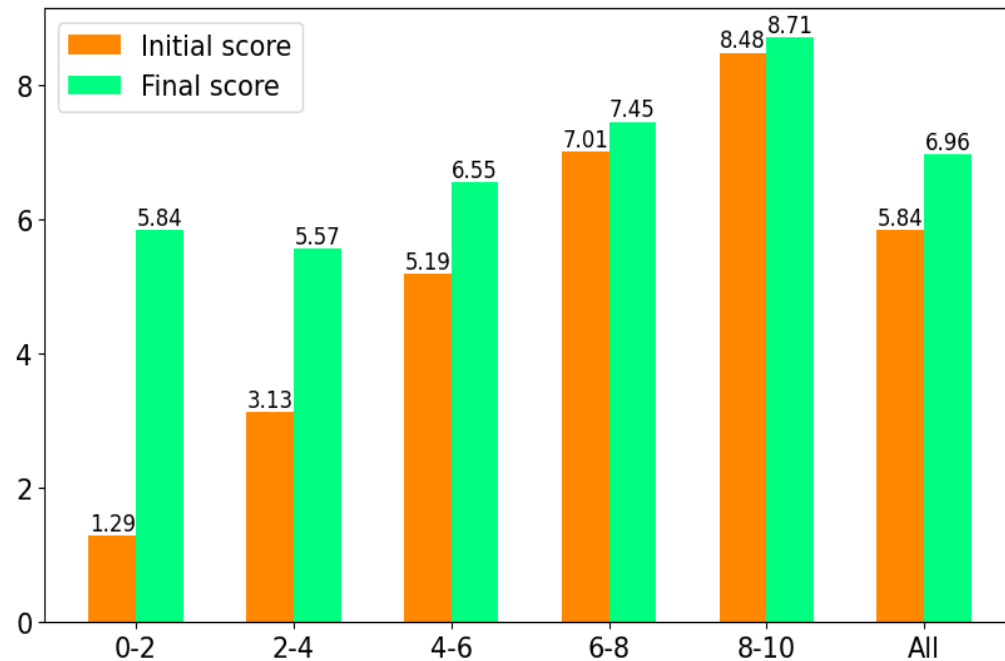
Examples of facilitating students to iteratively create high-quality scene descriptions with the automated feedback provided by the system.



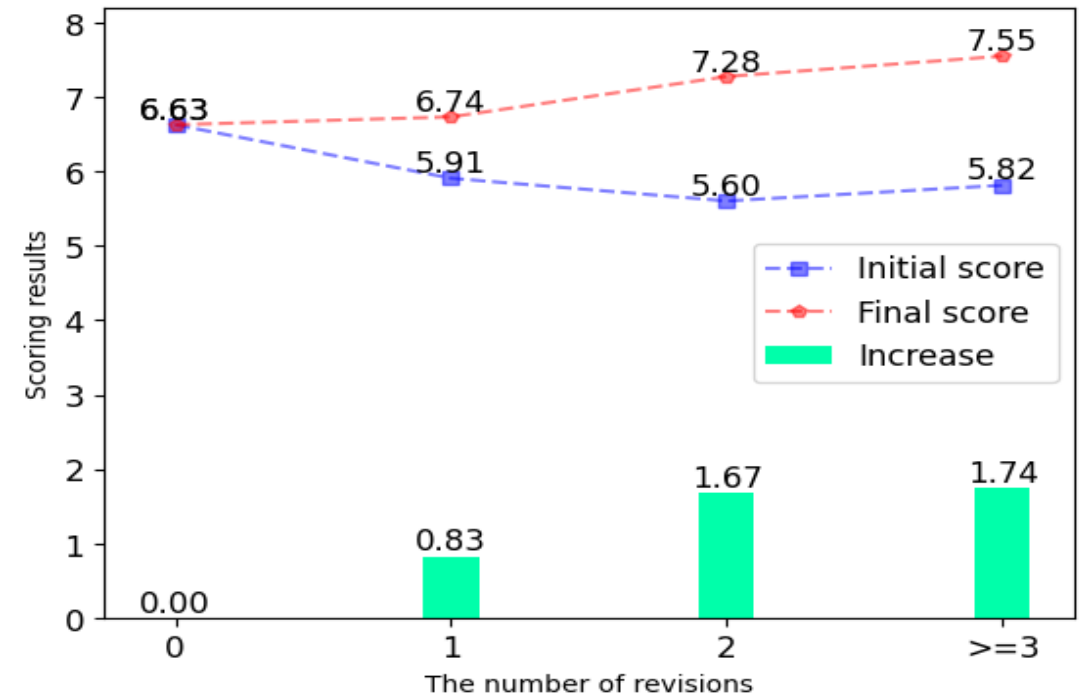
| Scenes | Descriptions being iteratively revised |                                                                                    | Scores on five dimensions |     |     |     |     | Total score |
|--------|----------------------------------------|------------------------------------------------------------------------------------|---------------------------|-----|-----|-----|-----|-------------|
|        | 1 <sup>st</sup>                        | Children                                                                           | 0.0                       | 0.0 | 0.5 | 0.2 | 0.2 | 0.9         |
|        | 2 <sup>nd</sup>                        | Children run in a snowy day                                                        | 2.0                       | 0.8 | 1.5 | 1.5 | 0.9 | 6.7         |
|        | 3 <sup>rd</sup>                        | Children in colourful clothes run in a snowy day                                   | 2.0                       | 1.7 | 1.5 | 1.7 | 0.9 | 7.8         |
|        | 4 <sup>th</sup>                        | In a snowy day, children in colourful clothes run at the school                    | 2.0                       | 1.8 | 1.5 | 2.0 | 0.9 | 8.2         |
|        | 1 <sup>st</sup>                        | There are three people fishing.                                                    | 0.9                       | 0.5 | 2.0 | 0.5 | 0.9 | 4.8         |
|        | 2 <sup>nd</sup>                        | Three people are fishing at the seaside.                                           | 1.7                       | 1.0 | 2.0 | 0.5 | 0.9 | 6.1         |
|        | 3 <sup>rd</sup>                        | An adult and two children are fishing at the seaside.                              | 3.0                       | 1.0 | 2.0 | 1.5 | 0.9 | 8.4         |
|        | 1 <sup>st</sup>                        | An old man was teasing a dog.                                                      | 2.0                       | 1.0 | 2.0 | 2.0 | 0.9 | 7.9         |
|        | 2 <sup>nd</sup>                        | An old man was playing with a dog on the lawn.                                     | 2.0                       | 1.5 | 2.0 | 1.5 | 0.9 | 7.9         |
|        | 3 <sup>rd</sup>                        | An old man wearing gray short sleeves was playing with a dog on the lawn.          | 2.0                       | 2.0 | 1.5 | 1.5 | 1.0 | 8.0         |
|        | 4 <sup>th</sup>                        | An old man wearing gray short sleeves was teasing a dog with branches on the lawn. | 2.5                       | 2.0 | 1.5 | 2.0 | 1.0 | 9.0         |

Zhao, R., Zhuang, Y., Xie Z. and Yu, P.L.H. (2024). Facilitating self-directed language learning in real-life scene description tasks with automated evaluation. *Computers & Education*, 2024: 105106.

# Feedback Help Students Progressively Improve Their Writing



Improvement trend of students' descriptions with the provided feedback



The impact of the number of revisions on improving the quality of students' descriptions



# Eye-Tracking Study: Students with Better Writing Spent More Time on Feedback

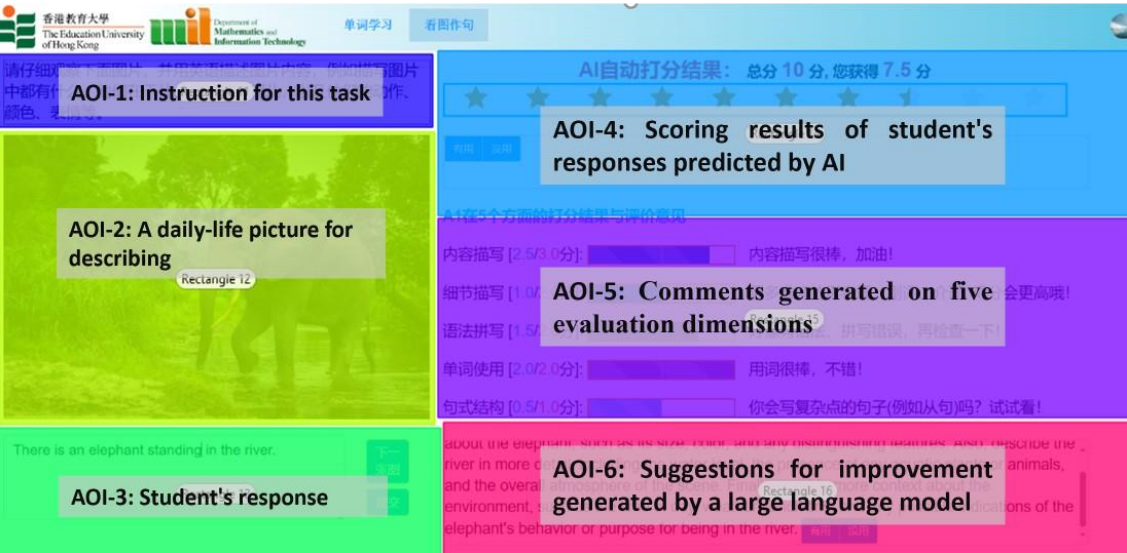


Table 1. Number of fixations and their durations on each AOI

|                                  | AOI-1<br>Instruction | AOI-2<br>Picture | AOI-3<br>Describing | AOI-4<br>Scoring | AOI-5<br>Comment | AOI-6<br>Suggestion |
|----------------------------------|----------------------|------------------|---------------------|------------------|------------------|---------------------|
| Fixation Duration_N              | 8.04                 | 100.12           | 368.35              | 23.64            | 56.02            | 152.47              |
| Fixation Duration_Mean (Seconds) | 0.13                 | 0.16             | 0.18                | 0.14             | 0.15             | 0.16                |
| Fixation Duration_Sum (Seconds)  | 1.19                 | 16.57            | 66.24               | 3.51             | 8.73             | 24.66               |

Students spent most time (24.66 sec) reviewing the suggestions for improvement, compared to the scoring results (3.51 sec) and comments (8.73 sec).

Zhao, R., Xie, Z., Zhuang, Y., Li, H. and Yu, P.L.H. (2024). *Enhancing Language Learning through Multimodal AI-Driven Feedback on Picture Descriptions: An Eye-Tracking Study*. In *Proceedings of the International Conference on Computers in Education (ICCE 2024)*, 25–29 November 2024, at Ateneo de Manila University, Quezon City, Philippines.

# Multimodal AI-Powered Language Enhancement (MAPLE)

Home

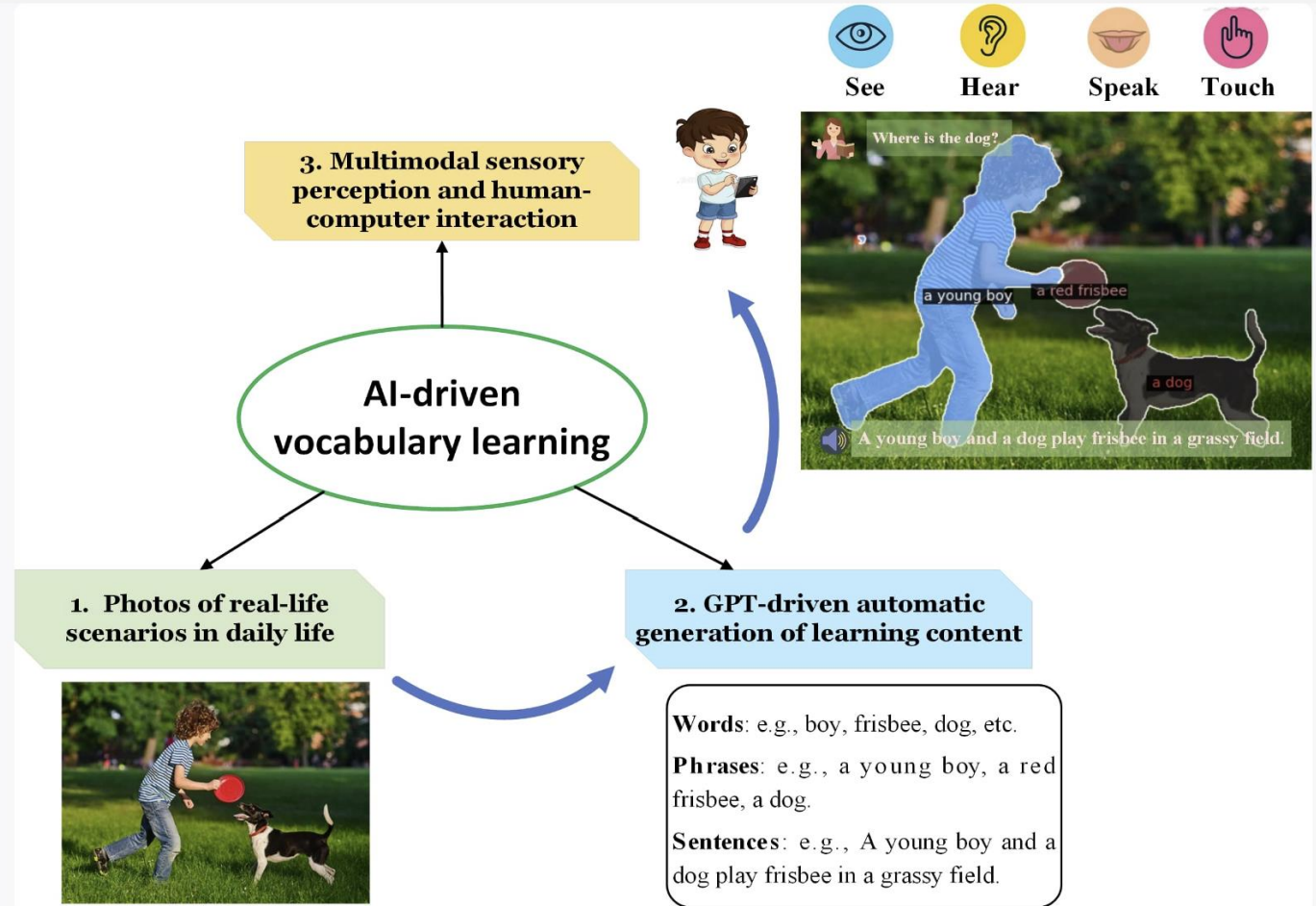
Let's Practice

Take a Test

Setting



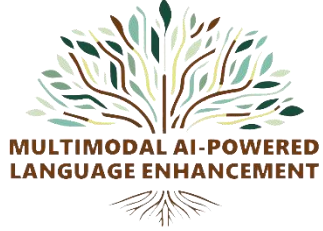
I am a multimodal AI model. Let's practice English together!



## Timeline

✓ You have already earned 44 scores in tasks!

i You have got 38 points.



# Multimodal AI-Powered Language Enhancement (MAPLE)



admin  
undefined undefined

Home

Let's Practice

Take a Test

Setting



admin  
undefined undefined

Home

Let's Practice

Take a Test

Setting



Public Tasks

Teacher's Tasks

Official Tasks



test4

Deadline 2024-11-30 00:00:00



test3

Deadline 2024-11-30 00:00:00



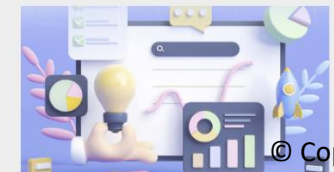
test2

Deadline 2024-11-30 00:00:00



test1

Deadline 2024-11-30 00:00:00



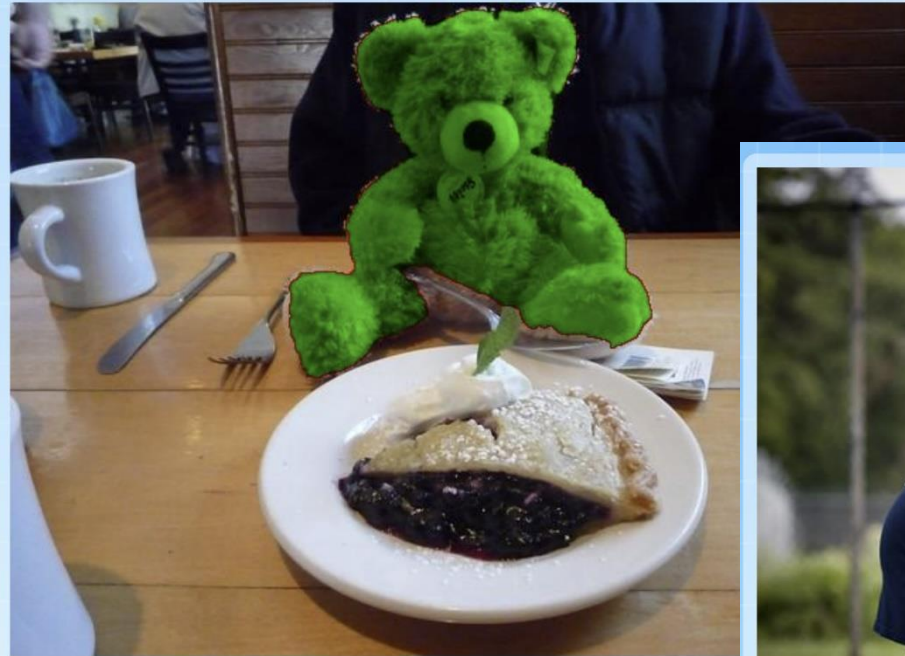


# MAPLE: Vocabulary Learning



dining table 🗣️

[Change an image](#)



teddy bear 🗣️

[Change an image](#)



tennis racket 🗣️

[Change an image](#)

# MAPLE: Vocabulary Learning



car 🔊



skateboard 🔊



road 🔊



# MAPLE: Practicing Language Skills Through Human-Computer Interaction



frisbee 🔊

[Change an image](#)

## Picture-cued language learning

You can learn vocabulary by clicking on objects in the image, and you can also do some exercises through interaction with the AI.

Give me another question



What activities are the dog and its owner involved in?

They are playing the frisbee.



In the image, there is a person and a dog engaged in an activity involving a frisbee. The dog appears to be jumping or preparing to jump up and catch the frisbee in its mouth. This is a common outdoor activity that many dog owners participate in with their pets to bond with them, provide exercise, and enjoy a fun, interactive game together.

Give me another question

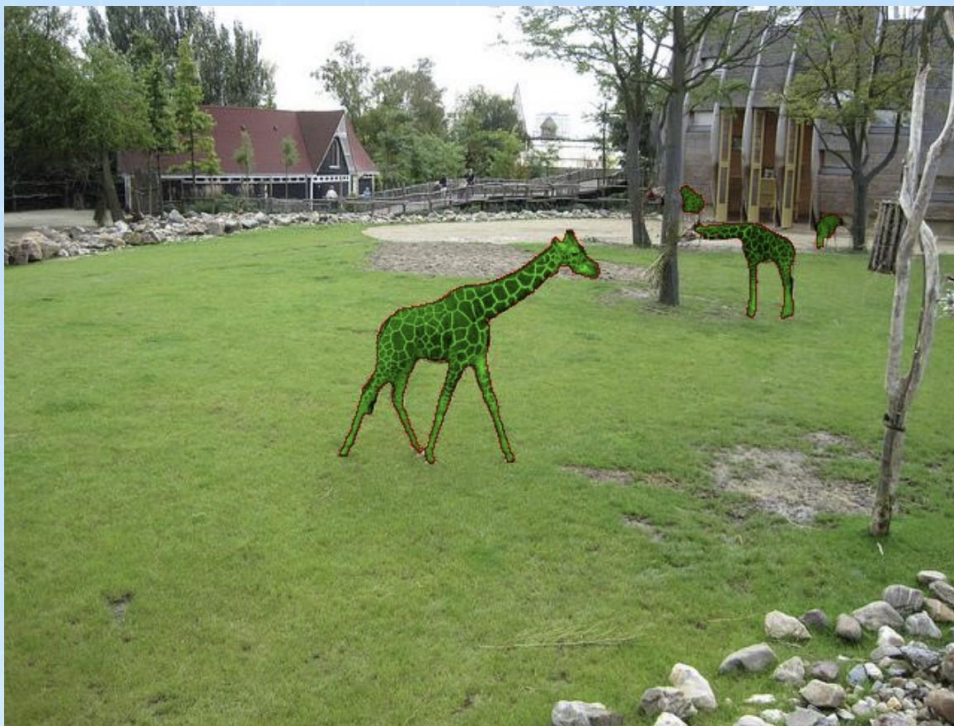


Sure, here's another question: How would you describe the relationship between the woman and her dog in this image?

Send



# MAPLE: Multiple Choices, Fill-In-The-Blanks & Describe a Picture



giraffe 🔊

New !



## Picture-cued language learning

You can learn vocabulary by clicking on objects in the image, and you can also do some exercises through interaction with the AI.

specialized enclosures that can accommodate their height and unique feeding habits. They may also be part of larger mixed-species habitats, showcasing a variety of wildlife from the same ecosystem or region where they naturally reside.



Now you can try...

[GAME] Multiple Choices

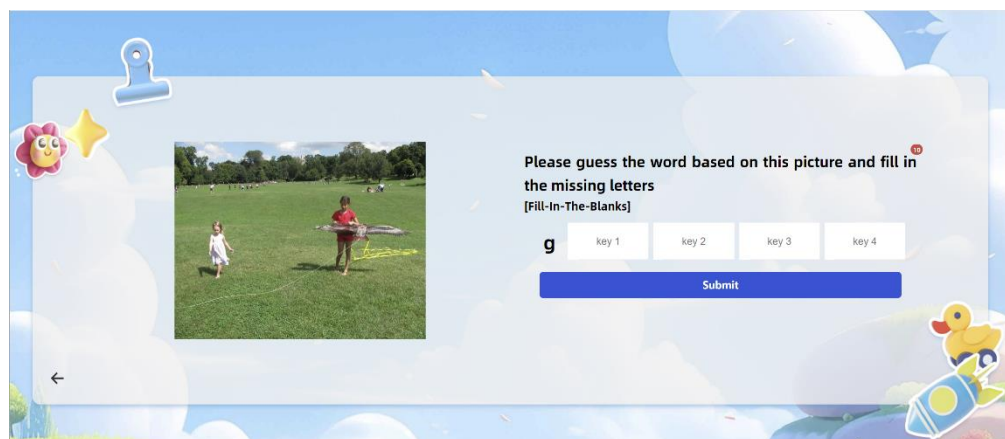
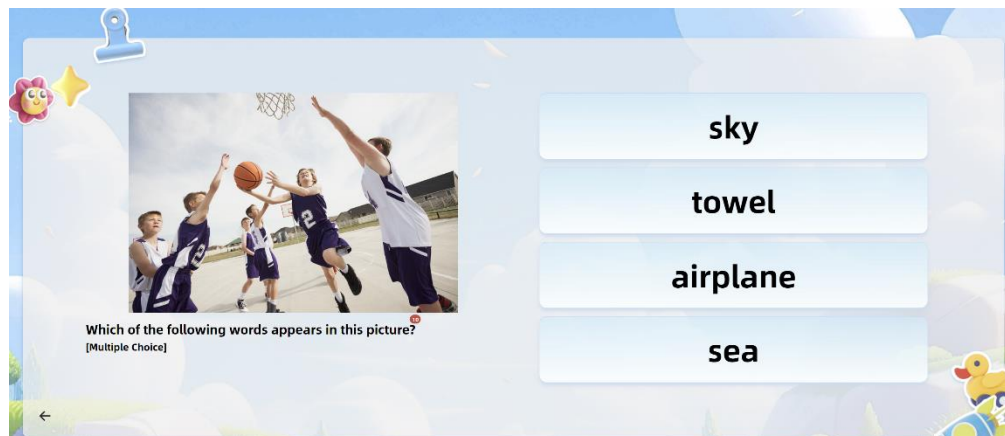
[GAME] Fill-In-The-Blanks

[Test] Describe a picture

One More Question

Send

# MAPLE: Multiple Choices, Fill-In-The-Blanks & Describe a Picture





Describe a picture

The woman is playing with a dog.

Submit

**Automated Scoring**

|               |         |      |
|---------------|---------|------|
| Comprehensive | 3 / 3   | 100% |
| Vividness     | 0.8 / 2 | 40%  |
| Mechanics     | 2 / 2   | 100% |
| Vocabulary    | 2 / 2   | 100% |
| Sentence      | 0.9 / 1 | 90%  |

Overall: 8.7 / 10

[Comments]

Overall comments: Your sentence does a good job of describing the main activity in the image, which is a woman and her dog engaging in play. You have captured the essence of the scene and made it clear what is happening. Comprehensiveness: Your sentence is concise yet informative. It covers the essential elements present in the picture, including the woman and the dog. To further enhance the descriptive aspect, you might consider adding more details about the environment or the specifics of their play activity. Vividness: While your sentence conveys the action effectively, to make it more vivid and engaging, you could use more sensory-rich language. For instance, you could describe how the dog is jumping up or how the woman appears focused on her task.

AI-generated feedback

# MAPLE: Teacher Resource Management

MAPLE **TEACHER**

- Home
- Schools Manage
- Teaching Resource
- Data
- Records

Tests Questions Questionnaires

Create New Question

| ID | Type               | Title                | Image                                                                                                                                                          | St |
|----|--------------------|----------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 2  | Describe a picture | [Describe a picture] | <br>sports ball<br>fence<br>playingfield<br>person<br>tennis racket         |    |
| 1  | Describe a picture | [Describe a picture] | <br>car road<br>grass car<br>sky<br>person<br>person<br>person<br>kite tree |    |

| Grades Classes Teachers Students |          |                     |         |                     |                     |                          |
|----------------------------------|----------|---------------------|---------|---------------------|---------------------|--------------------------|
| Add Batch Import                 |          | please select grade |         | please select class |                     | Q search...              |
| Student Id                       | Username | fullname            | Grade   | Class               | Create Time         | Actions                  |
| 50                               | stu50    | baron               | Grade 3 | 二班                  | 2024-07-14 10:37:40 | Data Edit Reset Password |
| 49                               | stu49    | charles             | Grade 3 | 二班                  | 2024-07-14 10:37:39 | Data Edit Reset Password |
| 48                               | stu48    | gordon              | Grade 3 | 二班                  | 2024-07-14 10:37:38 | Data Edit Reset Password |
| 47                               | stu47    | hale                | Grade 3 | 二班                  | 2024-07-14 10:37:37 | Data Edit Reset Password |
| 46                               | stu46    | carl                | Grade 3 | 二班                  | 2024-07-14 10:37:36 | Data Edit Reset Password |
| 45                               | stu45    | jack                | Grade 3 | 二班                  | 2024-07-14 10:37:35 | Data Edit Reset Password |
| 44                               | stu44    | bernie              | Grade 3 | 二班                  | 2024-07-14 10:37:33 | Data Edit Reset Password |
| 43                               | stu43    | archibald           | Grade 3 | 二班                  | 2024-07-14 10:37:32 | Data Edit Reset Password |

ENABLE

教大附中

xzavier

2024-11-15  
03:56:43

Copy and Create

6 items

10 / page

1



# MAPLE: Learning Records and Learning Logs

MAPLE **TEACHER**

admin

- Home
- Schools Manage
- Teaching Resource
- Data
- Records**

please select

Please Select Export Type

Record Id

Test Id

Student

Submit Time

Score

Records

4

5

教大附中 / 一班 /  
cornelia

2024-11-17 13:11:21

3

Pre-Questionnaire

Questions

Questionnaire

3

4

教大附中 / 一班 /  
cornelia

2024-11-17 13:10:23

20

Pre-Questionnaire

Questions

Questionnaire

2

3

教大附中 / 一班 /  
cornelia

2024-11-17 13:05:22

11

Pre-Questionnaire

Questions

Questionnaire

Pre-Questionnaire

Questions

Questionnaire

Result

ImageID: 2184

[Describe a picture]

Current Question Score

1.2

My Answer: 2222

Comprehensive

0.2 / 3

7%

Vividness

0 / 2

0%

Mechanics

0.6 / 2

30%

Vocabulary

0.2 / 2

10%

Sentence

0.2 / 1

20%

Overall: 1.2 / 10

Overall Comments: Your sentence, "2222," does not provide any descriptive information about the image. This makes it difficult for anyone reading your sentence to understand what is depicted in the picture. As a young student learning how to write, it's important to strive for clarity and detail in your descriptions so that your readers can engage with the images you are describing effectively. Comprehensiveness: The sentence lacks any specific details about the image content. It doesn't convey the main action or focus of the picture, which appears to be a family playing with a kite in a park. Vividness: Since your sentence is not descriptive, it lacks vividness.

Result

ImageID: 2181

[Fill-In-The-Blanks]

Current Question Score

10

My Answer: rass

Correct



ImageID: 2177

[Select the specified image element]

Current Question Score

0

My Answer:



Wrong



ImageID: 2175

[Multiple Choice]

Current Question Score

0

My Answer: towel

Wrong

# Extension to Immersive learning

Promoting immersive language learning by virtual reality and augmented reality

Real-life situations

Our AI models and learning support system

360° VR

Overlaying digital learning  
content onto real-life  
scenarios through AR



Virtual reality (VR) glasses



Augmented reality glasses



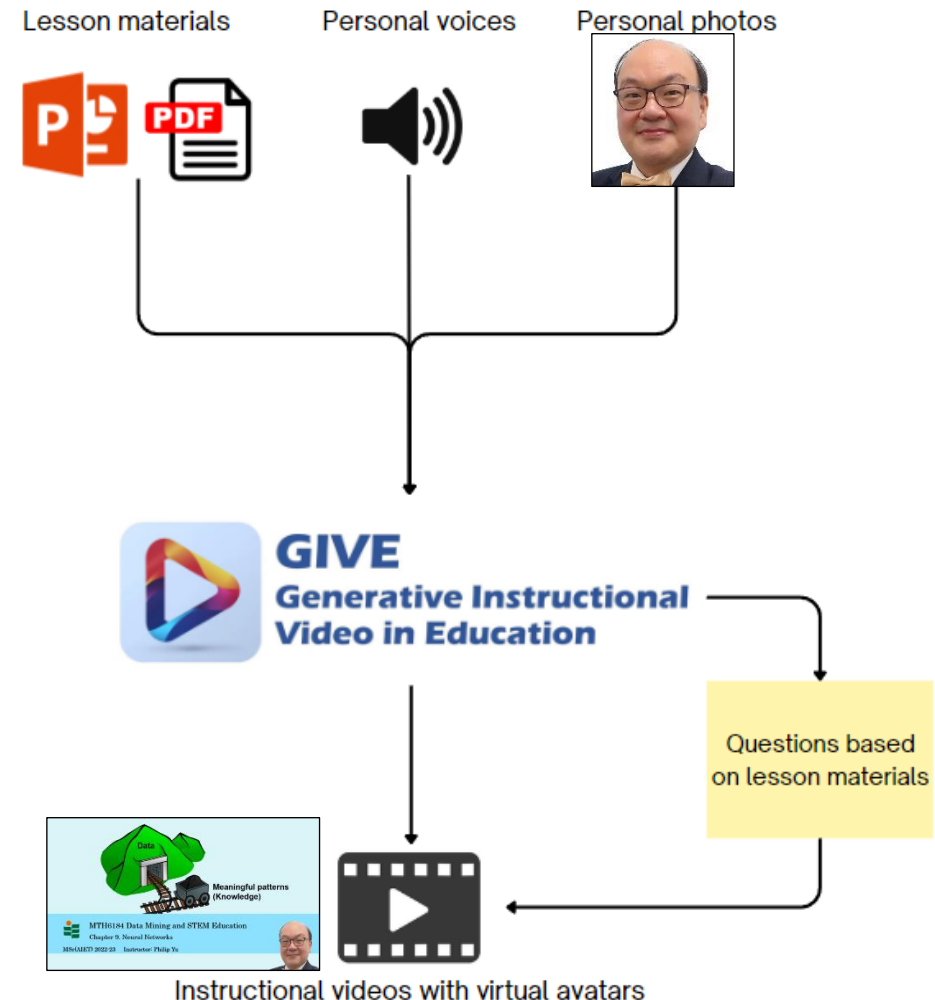
# Extension to Immersive learning





# Generative Instructional Video in Education (GIVE)

- Our latest **GIVE platform** proposes an innovative multimodal framework to generate customized instructional videos.
- The platform seamlessly integrates a **Large Language Model (LLM)**, a **Text-to-Speech model**, a **Talking Head Generation model** and other automated codes.
- It allows teachers to input their **lesson materials**, **personal audio chips**, and **personal photos** into the platform, aiming to automatically generate instructional videos with
  - a synthetic teacher's voice
  - precise lip-audio sync
  - natural talking head movementsand **questions** based on lesson materials.



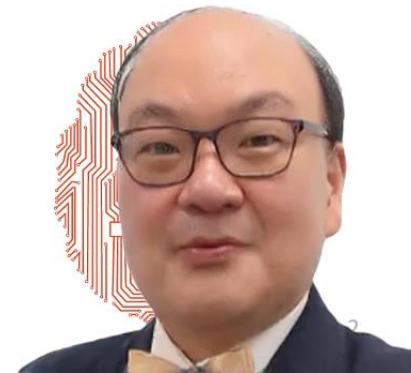
# Generative Instructional Video in Education (GIVE)

---

## Chapter Outcomes

After completing this chapter, you are able to:

- Understand the relation between a generalized linear model and a neural network
- Forecast a time series using neural networks



<https://tell.eduhk.hk/>



Multimodal AI Lab

[Products](#) ▾ [Activities](#) [Team](#) [Contact](#)

# The **Multimodal AI Solution** for Your Language Learning

An AI-powered automated scoring and feedback system for real-life scene description tasks that apply multimodal information to accurately assess students' language skills and provide immediate insights, enabling progressive improvement. The system incorporates level-specific adjustments, targeted at schools / teachers, which will solve the challenge of scalable and personalized language learning assessment. It has been successfully piloted and will set for launch in April 2024.

[Lets Get Started →](#)



# Join us and Take Part in our Experimental Study

---

Thank you for your attention. If you are interested and want to use our system, please scan the QR code and leave your contact information.



We expect that at least three classes in your school will participate, and they will use the system to learn English for more than eight weeks.

# Acknowledgment

## Funding Support:



## Research Team:



Dr Zhao Ruibin  
Post-doctoral Fellow



Mr Zhuang Yipeng  
PhD Student



Mr Zhang Tianyu  
PhD Student



Mr Xie Zhiwei  
Research Assistant



Mr Pan Haosen  
Research Assistant



Miss Zhang Jialin  
Research Assistant





香港教育大學

The Education University  
of Hong Kong



香港教育大學

The Education University  
of Hong Kong

# Thank You

Philip L.H. Yu

[plhyu@eduhk.hk](mailto:plhyu@eduhk.hk)

