

Cómo citar este documento:

Casas-Tost, Helena; González-Torrents, Isabel; Rovira-Esteva, Sara; Vargas-Urpí, Mireia (2025). *Herramienta de evaluación de actividades de comprensión escrita generadas con inteligencia artificial / Assessment tool for the evaluation of AI-generated reading comprehension activities*. URL: <https://ddd.uab.cat/record/313581>

[English translation provided below]

Índice / Table of contents

Herramienta de evaluación de actividades de comprensión oral generadas con inteligencia artificial.....	1
Evaluación de textos generados con IA	2
Evaluación de preguntas de comprensión lectora generadas con IA.....	3
Assessment tool for the evaluation of AI-generated reading comprehension activities .	4
Assessment of AI-generated texts	4
Assessment of AI-generated reading comprehension questions.....	5

Herramienta de evaluación de actividades de comprensión oral generadas con inteligencia artificial¹

Esta herramienta tiene el objetivo de proporcionar criterios para valorar la eficiencia y usabilidad de las actividades de comprensión lectora generadas con inteligencia artificial (IA). Está especialmente diseñada para la enseñanza de la lengua china como lengua adicional, pero es aplicable a otras lenguas.

¹ Esta herramienta se ha desarrollado en el marco del proyecto LT-LiDER para la investigación asociada al capítulo «Using GenAI in the teaching and learning of additional languages: Chinese as a case study» (Casas-Tost et al., en prensa). LT-LiDER (ref. 2023-1-ES01-KA220-HED-000161531) es un proyecto cofinanciado por la Comisión Europea.

Evaluación de textos generados con IA

Evalúa los siguientes aspectos en función del nivel de intervención que necesiten:

Nivel

- ☐ Vocabulario y/o gramática adecuados
- ☐ Vocabulario y/o gramática demasiado alto y no aceptable (ejemplos)
- ☐ Vocabulario y/ gramática demasiado alto pero aceptable (ejemplos)
- ☐ Vocabulario y/o gramática demasiado bajo (ejemplos)

«Aceptable» significa que los cambios necesarios son mínimos o que el texto incluye palabras que, aunque no se entiendan, se pueden deducir por el contexto.

Corrección lingüística (vocabulario, gramática y/o puntuación):

- ☐ Correcto
- ☐ Incorrecto, pero fácilmente corregible (ejemplos)
- ☐ Incorrecto y no aceptable por el volumen de incorrecciones (ejemplos)

Pragmática

- Aunque entiendas lo que dice el texto, ¿crees que hay elementos que un nativo no diría así? ¿Cuáles y cómo los dirías?
- Aunque entiendas lo que dice el texto, ¿crees que hay ideas que un nativo no incluiría en un texto así? ¿Cuáles?

Contenido:

- Diversidad y riqueza léxica: sí/no
- Contenido suficiente para evaluar la competencia (CE): sí/no

Evaluación global:

- ☐ Aceptable sin cambios
- ☐ Aceptable con cambios asumibles
- ☐ No aceptable

Tiempo total invertido en la corrección del texto:

Evaluación de preguntas de comprensión lectora generadas con IA

1. ¿Las preguntas se ajustan al nivel de lengua?

- ☐ Sí
- ☐ No, nivel demasiado alto
- ☐ No, nivel demasiado bajo

2. ¿Todas las preguntas son adecuadas?

- ☐ Sí
- ☐ No

En caso de que no, marca con colores las preguntas que no te parecen adecuadas. ¿Por qué no lo son? (marca todas las opciones que consideres)

- ☐ La respuesta es demasiado explícita en el texto (respuesta muy obvia y/o formulación idéntica en el texto). [amarillo]
- ☐ Los distractores en las opciones de respuesta son inverosímiles (no salen en el texto). [naranja]
- ☐ No son pertinentes para el texto porque no se relacionan con su contenido. [rojo]

3. Evaluación global:

- ☐ Aceptable sin cambios
- ☐ Aceptable con cambios mínimos/asumibles
- ☐ No aceptable

Assessment tool for the evaluation of AI-generated reading comprehension activities (Translation of the original questionnaire)

This tool aims to provide criteria for assessing the efficiency and usability of reading comprehension activities generated with artificial intelligence (AI). It is specially designed for the teaching of Chinese as an additional language, but is applicable to other languages.

Assessment of AI-generated texts

Level

- ☐ Appropriate vocabulary and/or grammar
- ☐ Vocabulary and/or grammar too advanced and not acceptable (examples)
- ☐ Vocabulary and/or grammar too advanced but acceptable (examples)
- ☐ Vocabulary and/or grammar too basic (examples)

"Acceptable" means that the necessary changes are minimal or that the text includes words that, while not immediately understood, can be inferred from context.

Linguistic Accuracy (vocabulary, grammar, and/or punctuation)

- ☐ Correct
- ☐ Incorrect but easily correctable (examples)
- ☐ Incorrect and not acceptable due to the number of errors (examples)

Pragmatics

- Even if you understand the text, do you think there are elements that a native speaker would not phrase this way? If so, which ones, and how would you rephrase them?
- Even if you understand the text, do you think there are ideas that a native speaker would not include in such a text? If so, which ones?

Content

- Lexical diversity and richness: Yes/No
- Sufficient content to assess reading comprehension (RC): Yes/No

Overall Evaluation

- ☐ Acceptable without changes
- ☐ Acceptable with minor revisions
- ☐ Not acceptable

Total time spent on text revision: _____

Assessment of AI-generated reading comprehension questions

1. Do the questions match the language level?

- ☐ Yes
- ☐ No, level too advanced
- ☐ No, level too basic

2. Are all the questions appropriate?

- ☐ Yes
- ☐ No

If not, highlight in colours the questions that you find inappropriate. Why are they not suitable? (Select all that apply)

- ☐ The answer is too explicit in the text (too obvious and/or identical wording as in the text). [yellow]
- ☐ The distractors in the answer choices are unrealistic (not present in the text). [orange]
- ☐ They are not relevant to the text because they are unrelated to its content. [red]

3. Overall evaluation:

- ☐ Acceptable without changes
- ☐ Acceptable with minor changes
- ☐ Not acceptable