ChatGPT's Performance and the Use of Argumentative Structures in Romanian Educational Contexts

Roxana Rogobete[1], Ana-Maria Radu-Pop[1]

[1]West University of Timisoara (Romania)

Summary

Introduction – Use of AI in Educational Contexts (Writing Tasks)

Study Design and Research Steps

Student Evaluation of ChatGPT's Academic Writing Performance

Pedagogical Implications: Integrating AI in Teaching and Academic Writing

Use of AI in Educational Contexts (Writing Tasks)

ChatGPT's Performance and the Use of Argumentative Structures in Romanian Educational Contexts,
Roxana Rogobete, Ana-Maria Radu-Pop,
Faculty of Letters, History, Philosophy and Theology,
West University of Timișoara

Use of AI in Educational Contexts - Romania (Writing Tasks)

- Tools like **ChatGPT,** Bard, LLaMA, and Claude – increasingly used

- Romanian Educational Landscape: conservative, formal academic writing standards

- **Argumentative writing**

- Central to national exams (e.g., National Evaluation, Baccalaureate, Definitivat, Titularizare = *Tenure Exam for Teachers, Final Teaching Qualification Exam*)

- More and more GenAI integrated in writing – ethical concerns? Plagiarism/originality

- integrating LLMs into writing instruction?

ChatGPT's Performance and the Use of Argumentative Structures in Romanian Educational Contexts,
Roxana Rogobete, Ana-Maria Radu-Pop, Faculty of Letters, History, Philosophy and Theology, West University of Timişoara

4

Academic Writing in Romania?

Academic Writing in Romania?

- **ROGER platform**, the Corpus of Romanian Academic Genres, undergraduate academic writing

- No dedicated corpus of pre-university student writing, work in progress – PhD theses for equivalent learner corpus, but none is currently available for secondary or high school writing

- Discrepancies in writing pedagogy and student performance across educational transitions

- Writing practices promoted in secondary education, which are often reproductive, prescriptive, and exam-oriented vs the expectations of university-level writing (requires analytical, discipline-specific, and rhetorically competent production, see Tucan et al. 2020)

- Need for enhanced pedagogical strategies and empirical resources that reflect the realities of student writing (how ChatGPT has a potential as a writing assistant)

ROGER corpus
a bilingual digital collection of student academic writing

2,050 texts · 8 disciplines · 20,325,396 characters · Search by word or phrase · Bilingual

About

The main characteristics of the ROGER corpus ("ROGER" as in: "Corpus of ROmanian academic GEnRes) are:

FLIFT

Study Design and Research Steps

Study Design and Research Steps

- MA students specializing in curriculum development for Romanian language and literature

- **3 teams**, each focusing on: National Evaluation (8th grade), Baccalaureate Exam (12th grade), Teacher Certification Exams (Definitivat & Titularizare)

- Data Collection – Essay Prompts: exam items (2014–2025) requiring argumentative/discursive essays

- Data Collection

  - Human-Written Responses

  - AI-Generated Responses

  - Hybrid Texts Creation (AI-enhanced)

ChatGPT's Performance and the Use of Argumentative Structures in Romanian Educational Contexts,
Roxana Rogobete, Ana-Maria Radu-Pop, Faculty of Letters, History, Philosophy and Theology, West University of Timişoara

8

Study Design and Research Steps

- Evaluation Phase – Blind Assessment – Group 1

- A new group of prospective teachers (Group 1) evaluated essays

- Given the official national rubric for their exam level

- Essays presented anonymously and randomly, No indication of origin: human, AI, or hybrid

- **Key Findings:** Human-written essays outperformed AI in 2 out of 3 exam levels; ChatGPT Limitations

- AI Improvement: Follow-up Activity - Use ChatGPT to improve writing

ChatGPT's Performance and the Use of Argumentative Structures in Romanian Educational Contexts,
Roxana Rogobete, Ana-Maria Radu-Pop, Faculty of Letters, History, Philosophy and Theology, West University of Timişoara

9

FLIFT

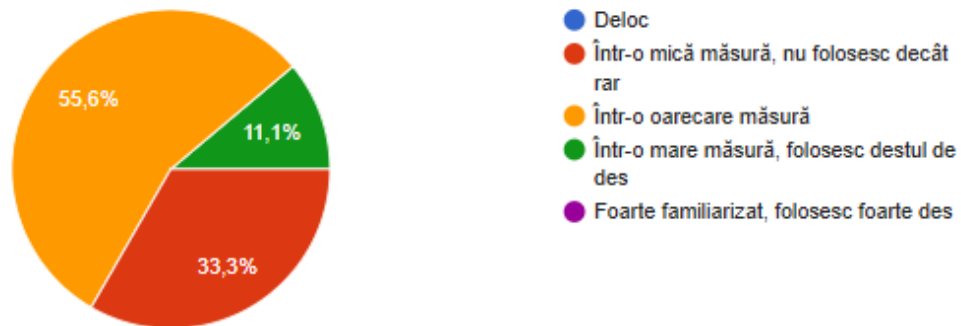# Study Design and Research Steps

- **Group 2: Validating Human-AI Collaboration**

- No prior involvement in the project, tasked with assessing:

  - Original ChatGPT essay

  - Original human-written essay

  - Enhanced human-AI collaborative version

- Evaluation: same national rubric

- **Key Results:** Enhanced (Version 3) essays scored highest across all prompts

- **Conclusion:** Guided AI use = effective pedagogical tool

ChatGPT's Performance and the Use of Argumentative Structures in Romanian Educational Contexts,
Roxana Rogobete, Ana-Maria Radu-Pop, Faculty of Letters, History, Philosophy and Theology, West University of Timişoara

10

Student Evaluation of ChatGPT's Academic Writing Performance

ChatGPT's Performance and the Use of Argumentative Structures in Romanian Educational Contexts,
Roxana Rogobete, Ana-Maria Radu-Pop, Faculty of Letters, History, Philosophy and Theology, West University of Timișoara

Student Evaluation of ChatGPT's Academic Writing Performance

Prior experience with ChatGPT:

- Deloc
- Într-o mică măsură, nu folosesc decât rar
- Într-o oarecare măsură
- Într-o mare măsură, folosesc destul de des
- Foarte familiarizat, folosesc foarte des

55,6%
11,1%
33,3%

Most students used ChatGPT for rephrasing, grammar checking, and improving text structure, with several also mentioning synonym search and idea generation.

Use was mostly limited and task-specific, showing reliance on AI for language polishing rather than content creation or deep analysis.

ChatGPT's Performance and the Use of Argumentative Structures in Romanian Educational Contexts,
Roxana Rogobete, Ana-Maria Radu-Pop, Faculty of Letters, History, Philosophy and Theology, West University of Timișoara

Evaluation - 5-point Likert scale

- ChatGPT responded effectively to the specific requirements of the academic writing tasks.
- ChatGPT helped me save time and effort in the writing process.
- Using ChatGPT contributed to the development of ideas in writing.
- The suggestions received through ChatGPT improved the overall quality of the texts.
- The suggestions received through ChatGPT improved the overall content of the texts.
- The suggestions received through ChatGPT improved the overall coherence of the texts.
- The suggestions received through ChatGPT improved the text structure.
- The suggestions received through ChatGPT improved the clarity of the texts.
- The suggestions received through ChatGPT created a more elaborate expression.
- I noticed an improvement in Romanian expression following the use of ChatGPT.
- Overall, the text "improved" with ChatGPT would receive a higher score in evaluation.
- I noticed a more elevated language style in texts following the use of ChatGPT.
- ChatGPT aligns well with the indicators in the rubric.

Evaluation - 5-point Likert scale

Time-saving benefits were acknowledged (mean 3.11), but confidence in writing improvement was low (mean 2.11).

Students doubted ChatGPT's ability to elevate Romanian writing or outperform human-written texts in formal assessments (both mean 2.0).

Difficulty distinguishing correct from incorrect AI suggestions (mean 3.22) highlights the need for critical digital literacy skills.

Strong support exists for integrating ethical AI use education at university (mean 4.0) and high school (mean 3.77) levels.

Findings reveal a pedagogical paradox: students value AI for accessibility and drafting but remain cautious about its educational accuracy and standard alignment.

Future AI implementation should include training on prompt engineering, rubric-based revisions, and ethical AI use to promote critical and informed engagement.

Student Reflections on Effective Prompting Strategies for ChatGPT

- Persona-based prompts (e.g., instructing ChatGPT to adopt an evaluator's perspective) often improved response coherence and structure

- Clear, specific instructions emphasizing natural, human-like language and consistent discourse style led to better AI outputs.

- Maintaining prompt continuity across multiple chats helped prevent confusion and improved task execution.

- Users' familiarity and skill with crafting prompts strongly influenced ChatGPT's responses.

- ChatGPT was seen as most effective for micro-level help—such as synonyms, reformulations, or technical language—rather than generating entire academic essays.

Limitations in ChatGPT's Output and the Need for Prompt Reformulation

- The tool frequently relied on formulaic, clichéd expressions despite requests for originality.

- Complex, multi-layered prompts often confused the model, resulting in incomplete or shallow elaboration.

- ChatGPT had difficulty avoiding redundancy and repetition, leading to circular arguments in some responses.

- Most students emphasized the need for critical human review to ensure authenticity, coherence, and appropriateness of AI-generated texts.

Pedagogical Implications: Integrating AI in Teaching and Academic Writing

Pedagogical Implications: Integrating AI in Teaching and Academic Writing

- AI as a cognitive scaffold vs Concerns about AI overuse

- Human-AI collaboration enhances academic writing

- Need for critical verification

- Pedagogical shift needed

- Call for AI literacy in education

ChatGPT's Performance and the Use of Argumentative Structures in Romanian Educational Contexts,
Roxana Rogobete, Ana-Maria Radu-Pop, Faculty of Letters, History, Philosophy and Theology, West University of Timișoara

References

Jiang, F., & Hyland, K. (2024). Does ChatGPT argue like students? Bundles in argumentative essays. *Applied Linguistics*. https://doi.org/10.1093/applin/amae052

Bašić, Ž., Banovac, A., Kružić, I., & Jerković, I. (2023). ChatGPT-3.5 as writing assistance in students' essays. Humanities and social sciences communications, 10(1), 1-5. https://www.nature.com/articles/s41599-023-02269-7.pdf

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. Assessing Writing, 57, 100752. https://doi.org/10.1016/j.asw.2023.100752

Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education*, *49*(5), 847-864.

Dhillon, P. S., Molaei, S., Li, J., Golub, M., Zheng, S., & Robert, L. P. (2024, May). Shaping human-ai collaboration: varied scaffolding levels in co-writing with language models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1-18).

Malik, A. R., Pratiwi, Y., Andajani, K., Numertayasa, I. W., Suharti, S., & Darwis, A. (2023). Exploring Artificial Intelligence in Academic Essay: Higher Education Student's Perspective, International Journal of Educational Research Open, Volume 5, 100296, ISSN 2666-3740, https://doi.org/10.1016/j.ijedro.2023.100296, https://www.sciencedirect.com/science/article/pii/S2666374023000717

Oravițan, A., Chitez, M., Bercuci, L., & Rogobete, R. (2022). Using the bilingual Corpus of Romanian Academic Genres (ROGER) platform to improve students' academic writing, in Intelligent CALL, granular systems and learner data: short papers from EUROCALL 2022, Research-publishing.net, p. 315-321.

Tucan, D., Rogobete, R., Radu-Pop, A.-M., & Chitez, M. (2020). Cât de pregătiți sunt elevii de liceu pentru scrierea academică de nivel universitar? Studiu didactic contrastiv bazat pe date de corpus lingvistic, in „Analele Universității de Vest din Timișoara. Seria Științe Filologice", 58, 2020, 69-92. https://analefilologie.uvt.ro/wp-content/uploads/2022/01/D-Tucan_R-Rogobete_M-Chitez_AM-R_Pop_Anale-Litere-2020.pdf

Crașovan, E., & Rogobete, R. (2020). Analiza de text literar: o perspectivă cantitativă asupra textelor studențești. n „Analele Universității de Vest din Timișoara. Seria Științe Filologice", 55, 2020, 93-114. https://codhus.projects.uvt.ro/wp-content/uploads/2020/12/E-Crasovan_R_Rogobete_Anale-Litere-2020.pdf

Nusivera, E., & Hikmat, A. (2025). Integration of Chat-GPT Usage in Language Learning Model to Improve Argumentation Skills, Complex Comprehension Skills, and Critical Thinking Skills. *IJLTER. ORG*, *24*(2), 375-390.