Developing an AI chatbot to improve coherence in academic writing in English for multilingual students

Laila Haddad-Najjar, Tzipora Rakedzon

Introduction

Writing is a highly demanding cognitive process that requires comprehending texts, extracting key ideas, and transforming them into new written discourse (Nunan, 2003). It encompasses a set of distinct, interrelated thinking processes that writers actively coordinate while composing, such as planning, evaluating, and making decisions based on an evolving network of goals, through which writing proficiency develops (Flower & Hayes, 1981). These challenges are especially evident among EFL learners due to underdeveloped linguistic skills, limited vocabulary, and less background knowledge (Grabe & Zhang, 2013; Siekmann et al., 2022; Teng, 2020). Thus, L2 writers benefit from structured scaffolding, meaningful tasks, feedback, and opportunities for extended practice in academic genres (Grabe and Zhang, 2013).

While writing in EFL is demanding for the learners, academic writing poses an even greater challenge. Academic writing is formal written work produced in an academic setting, intended to inform, analyze, or argue based on evidence and research (Nunan & Choi, 2023). According to Johnson (2016), academic writing is an objective and goal-oriented process of locating sources, collecting data, planning, generating ideas, rereading, and revising. It typically starts with gathering information and then progresses to a more complex stage of organizing it into a structured form (Teng et al, 2019; Teng 2020). In academic writing, and writing in general, well-organized arguments that are relevant and easy to follow bring about coherent writing (Nunan & Choi, 2023).

Coherence and cohesion, key challenges in academic writing, are elucidated by Halliday and Hasan (1976; 2014) in their seminal work which defines coherence in relation to cohesion. While cohesion denotes the explicit linguistic links, coherence reflects how these links combine to create a logically connected and interpretable whole. They view coherence as the overall unity and meaningful flow of a text that arises from the reader's ability to interpret cohesive ties. In a subsequent contribution, Johns (1986) argues that coherence involves both text-based and reader-based dimensions, as it depends on how propositions are organized and connected. Crossley et al. (2016) distinguish cohesion as a set of measurable textual features from coherence, which refers to readers' mental

representations of text meaning. As such, semantic cohesion measures, such as Latent Semantic Analysis (LSA), reflect conceptual overlap beyond surface lexical ties and are therefore commonly interpreted as computational indices of global coherence.

Academic writing is challenging for EFL learners, and achieving coherence often emerges as one of the most difficult aspects. Silva's (1993) review of studies with advanced EFL tertiary students shows that their texts tend to have simpler structures with missing parts, including key ideas, and exhibit weaker coherence omission of key ideas. Hinkel (2011) argues that difficulties in achieving coherence often lead L2 learners to produce simplified text structures and to struggle with fully topic development and thesis statements.

Given these constraints, contemporary AI tools provide substantial support for improving language accuracy and transforming how learners develop academic writing, especially in EFL contexts (Fathi & Rahimi, 2024). Yuan et al (2024) found ChatGPT supported EFL learners by scaffolding literacy development, particularly in writing fluency, content generation, and subject-matter understanding. Language support and surface-level accuracy are also strengthened in academic writing, as AI systems improve grammar and vocabulary (Crompton et al., 2024). Jiang and Hyland (2025), in comparing essays generated by an AI generator (ChatGPT) to essays written by university students, found that while ChatGPT produces clear and coherent essays due to its algorithmic training, student writing varies because of individual and instructional differences.

The question of whether AI tools effectively address semantical cohesion as a proxy for coherence in academic writing remains notably under-explored, and the limited existing studies on AI and coherence report mixed findings: one chatbot acted as a "digital supervisor," offering formative feedback to improve coherence in dissertation synopses (Krumsvik,, 2024), while others found AI feedback limited to surface-level features, with little impact on coherence in EFL writing (Yoon et al., 2024). In comparing automated coherence metrics LLMs to human judgement, Chhun et al., 2024 reported that AI frequently fails to maintain coherence throughout long written texts, whereas Cohen et al. (2025) contend that LLM-based measures can successfully identify gaps in connectedness in scientific texts. Similarly, Morris et al. (2025) found that configured LLMs were more sensitive than human raters to linguistic features when evaluating coherence indices.

While AI offers a wide range of affordances for academic writing, the role AI plays in supporting coherence in students' academic writing is still inconclusive. As such, concerns remain about the depth of AI-generated feedback, which often focuses on surface-level features, such as transitions and repetition. This suggests that continued refinement of AI models could enhance their

effectiveness in fostering coherence in academic writing. Not only does the impact of AI use on coherence remain underexplored, but also the way speakers of typologically different languages leverage AI tools in academic writing in English remains unclear. To address this gap, we compare the works of Hebrew and Arabic- speaking students enrolled in academic writing courses and explore the role of AI in enhancing academic writing coherence and their attitudes towards the AI use.

Aims and research questions

Our study aims to examine the role of AI in enhancing academic writing coherence, with focus on semantic cohesion, and to explore how students from diverse linguistic backgrounds perceive and engage with AI in writing. To this end, it is guided by two research questions: (1) To what extent does AI use improve coherence in academic writing, and how does this differ between Hebrew- and Arabic-speaking students? (2) To what extent does semantic cohesion, as measured by LSA, change in L2 learners' writing from pre-test to post-test following AI intervention? And (3) How do students from each linguistic group engage with the chatbot to improve coherence, and how do they perceive its role?

Methods

Participants

Convenience sampling is used in this study: participants include PhD students from different STEM departments at the Technion, enrolled in compulsory academic writing courses during the 2025 Winter Semester. Participants in these courses come from a variety of linguistic backgrounds. We will compare Hebrew and Arabic speakers, who represent the majority and minority populations in Israel. Both languages, as Semitic, differ markedly from English, which may pose additional coherence challenges in English academic writing. Participants from two course groups are assigned to two groups, an experimental group and a control group.

Procedure

As part of the academic writing course, students bring a draft of an introduction they are currently working on. They participate in an interactive lesson (lecture + in-class writing tasks) on the flow and coherence of academic texts. Students are then asked to complete the task on coherence by using a customized ChatGPT developed based on global cohesion indices (Crossley, 2016). The chatbot takes students through 5 guiding questions about logical paragraph structure and indices related to repeated key words, verb synonyms, noun synonyms, pronouns referring to previously

mentioned ideas, and connectors. In response to each question, students indicate where these cohesive devices appear in their drafts or where they intend to incorporate them.

The experimental group receives AI feedback on their responses, without rewriting their original text, whereas the control group completes the same writing tasks guided by the same questions but without AI feedback. All students use the chatbot via the De-Jargonizer platform (Rakedzon et al., 2017), a public, freely accessible tool for AI-based science communication training. After receiving feedback, students revise their texts accordingly and submit a post-task summary at the end of the intervention.

Data Analysis

Perceptions and experiences are explored through questionnaires and focus groups. Textual changes in coherence will be analyzed using global cohesion indices via TAACO (Crossley et al., 2016) with pre–post differences assessed via paired samples t-tests. Qualitative data from questionnaires and focus groups will undergo thematic analysis.

Institutional Description:

The Technion–IIT is a STEM (Science, technology, engineering, and math) research university with over 18 departments. It has approximately 15,000 students, including roughly 10,000 undergraduates and 5,000 graduate students. Students come from diverse backgrounds, including Jews, Arabs, and around 1,000 international students. Arab students make up 20% of the Technion's student population, equal to their percentage in Israel's overall population. According to recent data, the Technion has a high percentage of women: over 40% of graduate students, 48% of undergraduates, and approximately 50% of Arab undergraduate students are women.

The Technion is a global hub for innovation, science, and engineering research, as well as strong industry partnerships that drive technology. Since science and technology rely on communication, from understanding existing work to presenting new ideas, the Department of Humanities and Arts helps students develop these essential capabilities. Its Academic English and Academic Writing program trains STEM students, both undergraduate and graduate, to communicate their research in conferences and publish in peer-reviewed journals. Today, the department remains committed to developing AI-enhanced writing pedagogy to stay current with emerging trends in AI and writing.

Key Theorists:

Flower and Hayes's (1981) Cognitive Process Model of Writing conceptualizes writing as a recursive, goal-driven problem-solving activity involving the interaction of the task environment, the writer's long-term memory, and the writing process. Coherence is embedded in these processes as writers generate ideas, organize them, and monitor their evolving text. This model informs our study by explaining how AI tools may help improve coherence, functioning as an external scaffold that reduces cognitive load and enhances writers' ability to plan and revise effectively.

Vygotsky's (1978) Sociocultural Theory emphasizes that culture provides the linguistic systems, symbols, and intellectual tools that shape how learners think, communicate, and solve problems. The language, norms, and interaction styles with which learners grow up shape the guidance they receive and how they process information: this is mediated by a 'More Knowledgeable Other' (e.g., teacher, parent) through cultural interactions within the Zone of Proximal Development (ZPD). Technology can also function as a More Knowledgeable Other, and thus, this theory informs our study. We attempt to leverage AI-based guidance through coherence indices as a scaffolding tool that enables learners from diverse linguistic and cultural backgrounds to improve their writing outcomes.

Halliday and Hasan's (1976; 2014) cohesion taxonomy provides a foundational framework for describing how linguistic elements create unity and continuity in a text. They identify major cohesion types, including reference through linking words/phrases/conjunctions; substitution; ellipsis; and lexical cohesion. Together, these categories explain how texts achieve connectedness and coherence beyond just grammar.

Glossary

- **Academic writing**: writing produced to communicate research or scholarly ideas to academic audiences (within or across disciplines) under the conventions of formal publication.
- **Coherence**: the overall logical unity and meaningful flow of a text, enabling the reader to interpret how ideas relate across sentences, paragraphs, and the whole text.
- **Cohesion**: the network of explicit linguistic links (e.g., reference, conjunctions) that bind sentences and paragraphs together to support the reader's comprehension.
  - **Local cohesion**: refers to sentence-level links

- o **Global cohesion:** refers to connections across larger units such as paragraphs and aligns closely with coherence.
  - o **Tool for the Automatic Analysis of Text Cohesion** (**TAACO**): a digital tool to assess cohesion.
- **Semantic overlap:** refers to the degree of conceptual similarity across sentences or larger text units and is commonly operationalized using LSA-based measures.
- **AI-based learning / AI-enhanced writing pedagogy**: instructional approaches that integrate AI tools to support the writing-learning process (e.g., feedback, revision, tutoring), while excluding AI use for full text generation.
- **Hebrew- vs. Arabic-speaking students**: in our university context, two main L1 populations — Hebrew speakers and Arabic speakers — for whom English often functions as a second or third language.
- **STEM students:** Science, technology, engineering, and math students.
- **Tool for the Automatic Analysis of Cohesion (TAACO) (Crossley et al. 2016)** – a digital tool to assess cohesion. The computational framework is based on Halliday & Hasan's cohesion taxonomy as a theoretical foundation for identifying and operationalizing cohesion and coherence. TAACO enables the analysis of cohesion-related measures such as lexical overlap, semantic similarity, and lexical diversity.
- **Latent Semantic Analysis (LSA; Landauer & Dumais, 1997**) is a corpus-based computational model, which conceptualizes meaning based on analyzing how words and passages tend to appear together. By representing words and texts in a reduced semantic space, LSA estimates semantic similarity and conceptual overlap in words/texts. In applied linguistics and writing research, LSA has been widely used to analyze semantic cohesion and coherence. (e.g., McNamara et al., 2014).

References

Chhun, C., Suchanek, F. M., & Clavel, C. (2024). Do language models enjoy their own stories? prompting large language models for automatic story evaluation. *Transactions of the Association for Computational Linguistics*, *12*, 1122-1142.

Cohen, T., Xu, W., Guo, Y., Pakhomov, S., & Leroy, G. (2025). Coherence and comprehensibility: Large language models predict lay understanding of health-related content. *Journal of Biomedical Informatics*, *161*. https://doi.org/10.1016/j.jbi.2024.104758

Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13460

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior research methods*, *48*(4), 1227-1237.

Crossley, S. A., Kyle, K., & Dascalu, M. (2019). The Tool for the Automatic Analysis of Cohesion 2.0: Integrating semantic similarity and text overlap. *Behavior Research Methods*, *51*(1), 14–27. https://doi.org/10.3758/s13428-018-1142-4

Fathi, J., & Rahimi, M. (2024). Utilising artificial intelligence-enhanced writing mediation to develop academic writing skills in EFL learners: a qualitative study. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2024.2374772

Flower, L., & Hayes, J. R. (1981). *A Cognitive Process Theory of Writing* (Vol. 32, Issue 4).

Grabe, W., & Zhang, C. (2013). Reading and Writing Together: A Critical Component of English for Academic Purposes Teaching and Learning. *TESOL Journal*, *4*(1), 9–24. https://doi.org/10.1002/tesj.65

Halliday, M. A. K., & Hasan, R. (1974; 2014). *Cohesion in English*. Routledge.

Hinkel, E. (2011). *What research on second language writing tells us and what it doesn't*. In Handbook of research in second language teaching and learning (pp. 523-538). Routledge.

Johns, A. M. (1986). Coherence and academic writing: Some definitions and suggestions for teaching. *Tesol Quarterly*, *20*(2), 247-265.

Johnson, A. P. (2016). *Academic writing: Process and product*. Bloomsbury Publishing USA

Krumsvik, R. J. (2024). Chatbots and academic writing for doctoral students. *Education and Information Technologies*. https://doi.org/10.1007/s10639-024-13177-x

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*, *104*(2), 211.

Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (Eds.). (2007). *Handbook of latent semantic analysis*. Psychology Press.

Morris, W., Holmes, L., Choi, J. S., & Crossley, S. (2025). *Uncovering Differential Sensitivity Toward Linguistic Features of Cohesion in Large Language Models* (pp. 227–234). https://doi.org/10.1007/978-3-031-98465-5_29

McNamara, D. S., Crossley, S. A., & Roscoe, R. (2013). Natural language processing in an intelligent writing strategy tutoring system. *Behavior Research Methods*, 45(2), 499–515. https://doi.org/10.3758/s13428-012-0258-1

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press

Nunan, David., & Choi, Julie. (2023). *Clarity and coherence in academic writing: using language as a resource*. Routledge.

Siekmann, L., Parr, J. M., & Busse, V. (2022). Structure and coherence as challenges in composition: A study of assessing less proficient EFL writers' text quality. *Assessing Writing*, *54*. https://doi.org/10.1016/j.asw.2022.100672

Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. *TESOL quarterly*, 27(4), 657-677.

Teng, F. (2020). The role of metacognitive knowledge and regulation in mediating university EFL learners' writing performance. *Innovation in Language Learning and Teaching*, *14*(5), 436–450. https://doi.org/10.1080/17501229.2019.1615493

Teng, (Mark) Feng, & Huang, J. (2019). Predictive Effects of Writing Strategies for Self-Regulated Learning on Secondary School Learners' EFL Writing Proficiency. In *TESOL Quarterly* (Vol. 53, Issue 1, pp. 232–247). Wiley Blackwell. https://doi.org/10.1002/tesq.462

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological pro*cesses (Vol. 86). Harvard University Press.

Yoon, S.-Y., Miszoglad, E., & Pierce, L. R. (2023). *Evaluation of ChatGPT Feedback on ELL Writers' Coherence and Cohesion*. http://arxiv.org/abs/2310.06505

Yuan, Y., Li, H., & Sawaengdist, A. (2024). The impact of ChatGPT on learners in English academic writing: opportunities and challenges in education. *Language Learning in Higher Education*, *14*(1), 41–56. https://doi.org/10.1515/cercles-2023-0006