Check for updates

Research Article

Grammatical and Lexical Cohesion in Indonesian EFL Argumentative Essays

Septiana Wandira[1]* , La Ode Achmad Suherman[2], Dwi Nur Hadiyansah W.S.[3]
[1]Universitas Islam An Nur Lampung, Lampung, Indonesia
[2]Universitas Muhammadiyah Buton, Baubau, Indonesia
[3]Universitas Ibrahimi Situbondo, Situbondo, Indonesia
*septianawandira29@gmail.com

ABSTRACT

This study investigates grammatical and lexical cohesion in Indonesian EFL undergraduates' argumentative essays and examines how cohesion relates to writing quality. A small learner corpus of 15 timed essays was manually coded for reference, articles, conjunctions, and lexical reiteration following Halliday and Hasan's framework, and rated with an analytic rubric. Descriptive statistics and Spearman correlations were complemented by qualitative analysis of higher- and lower-rated texts. Quantitative results show frequent use of articles, basic conjunctions, and lexical repetition, but cohesion indices display weak, sometimes negative, associations with overall and coherence scores. Qualitative findings reveal that stronger essays are characterised by stable reference chains, a wider range of logical connectives, and purposeful lexical reiteration, whereas weaker texts rely on broken chains and mechanical repetition. The study argues that cohesion quality, rather than quantity, is more salient for raters and discusses implications for genre-based EFL writing instruction.

Keywords: Argumentative Writing; Cohesion; Coherence; EFL Learners; Indonesian University Students.

1. Introduction

Argumentative writing is a central genre in higher education because it requires students to articulate a clear position, support it with evidence, and anticipate counterarguments in a logically organised way. Beyond serving as a vehicle for assessment, it plays a crucial role in cultivating critical thinking and problem-solving skills (Ferretti & Graham, 2019; Kuhn, 2019). Recent work also shows how digital tools can be orchestrated to support argument construction: Sandra et al. (2024), for example, demonstrate that digital annotations, conversational agents, and collaborative concept maps can help undergraduates craft more compelling arguments. From an English for Academic Purposes perspective, Yasuda (2023) argues that learning to construct an argument means managing both general academic demands and discipline-specific expectations, making argumentative writing a key locus of academic socialisation.

For learners of English as a foreign language (EFL), however, producing effective argumentative essays is particularly demanding. It involves not only controlling grammar and lexis, but also organising ideas into coherent texts and aligning with genre-specific conventions. Dornbrack and Dixon (2014) note that even with explicit instruction, novice writers struggle to coordinate process and genre demands in argumentative tasks. Studies in EFL and ESL contexts report recurring problems with maintaining a clear

thesis, structuring arguments logically, and providing adequate support for claims (Abdollahzadeh et al., 2017; Mallahi, 2024). Recent work using sentence-reordering algorithms illustrates how fragile macro-organisation can be in EFL essays and how much improvement is possible when the logical sequence of sentences is optimised (Putra, Teufel, & Tokunaga, 2023). Research on complexity, accuracy, and fluency (CAF) also documents substantial variability in argumentative writing across learner groups and proficiency levels (Barrot & Gabinete, 2021). Genre-based pedagogies can provide useful scaffolding and foster more positive attitudes, yet many learners still plan and revise only minimally, leading to essays that are rhetorically underdeveloped (Bejarano & Chapetón, 2013; Hyland, 1990; Mahfoudhi, 2003).

Within this broader challenge, cohesion and coherence are widely recognised as key dimensions of writing quality because they allow readers to follow how ideas are linked and how arguments unfold. Building on Halliday and Hasan's model, cohesion can be understood as the network of grammatical and lexical ties such as reference, conjunction, and lexical reiteration that connect clauses and sentences on the surface of the text. Coherence, in contrast, is often conceived as the reader's experience of a text as meaningful and well organised (Faigley & Witte, 1981; Colomb & Griffin, 2004). Empirical research shows that coherent texts tend to receive higher quality ratings than texts that are merely grammatically accurate (McCulley, 1985; McNamara, Crossley, & McCarthy, 2010), and teaching studies suggest that explicit work on cohesive linkers and logical patterns can enhance readability and persuasive effect (Palmer, 1999; Zaheer, Rahman, & Sharma, 2024). Lexical cohesion in particular, achieved through repetition, synonyms, and semantically related words, plays an important role in building thematic chains and signalling relationships between propositions (Stotsky, 1983; Mojica, 2008; Kuo, 1995).

Research on cohesion in L2 writing, however, reveals a complex and uneven developmental picture. Learners typically rely on a limited repertoire of cohesive devices, especially basic conjunctions and simple lexical repetition, and only gradually diversify their cohesive resources (Neary-Sundquist, 2013; Bacha, 2002; Zanardi, 1994; Crowhurst, 1987). Longitudinal studies show that global cohesion develops slowly and non-linearly, with considerable individual variation in how cohesive practices evolve over time (Zhang & Zhang, 2024; Tabari & Wind, 2023; Zhang, 2023). Lower-proficiency learners may even use cohesive devices more frequently than higher-proficiency peers, but with less syntactic accuracy and less effective rhetorical control (Huang & Watzinger-Tharp, 2023). At the same time, many studies have focused primarily on the frequency and distribution of cohesive devices, paying less systematic attention to how cohesion relates to human judgements of writing quality. Classic critiques caution against treating cohesion counts as direct indices of coherence (Tierney & Mosenthal, 1983; Faigley & Witte, 1981), and more recent analyses likewise show that cohesion indices often have limited power to discriminate between high- and low-quality essays or to predict rubric-based scores (McNamara et al., 2010; Spiegel & Fitzgerald, 1990; Venegas, Castro-Cano, & Cornejo, 2024). Johnson (1992) similarly finds that raters attend more to overall coherence in content, organisation, and style than to the sheer number of cohesive ties.

These gaps are particularly salient in the Indonesian EFL context, where argumentative writing is a core outcome of university English programmes but empirical evidence on students' cohesive practices is still limited. Cross-cultural rhetorical studies show that Indonesian undergraduates tend to favour deductive patterns and may structure introductions and conclusions differently from native English speakers, with consequences for cohesion and argument development (Arsyad, 1999; Rhetorical pattern of the Indonesian EFL undergraduate students' writings, 2022). Corpus analyses of lexical bundles in Indonesian EFL argumentative essays report a predominance of

research-oriented bundles and a lack of structuring signals, suggesting a restricted range of cohesive resources used to organise texts (Oktavianti & Prayogi, 2022). Research on argument–counterargument structures indicates that such dialogic organisation is underutilised, partly due to proficiency constraints and limited exposure to genre-based instruction (Rusfandi, 2015). To date, however, few studies in this context have jointly examined grammatical and lexical cohesion, analytic ratings of writing quality, and close discourse analysis in students' English argumentative essays.

To address these gaps, the present study investigates the grammatical and lexical cohesion of Indonesian EFL students' argumentative essays and explores how cohesion patterns are associated with analytic ratings of writing quality. Adopting a corpus-informed mixed-methods design, the study first analyses the distribution and density of selected cohesive devices across a small learner corpus and then examines their relationships with overall and trait-based scores on an analytic rubric. In the second phase, a subset of higher- and lower-rated essays is subjected to qualitative discourse analysis grounded in Halliday and Hasan's framework, focusing on reference chains, conjunctive relations, and lexical reiteration. The study is guided by two research questions: (1) What patterns of grammatical and lexical cohesion characterise these Indonesian EFL learners' argumentative essays? and (2) How do these cohesion patterns relate to analytic ratings of writing quality and to qualitative differences between stronger and weaker texts? By integrating quantitative and qualitative perspectives in an under-researched context, the study seeks to refine our understanding of the often-contested relationship between cohesion and writing quality and to offer pedagogical implications for teaching cohesion as a functional resource in EFL argumentative writing.

2. Method

2.1 Research Design and Context

This study employed a corpus-informed mixed-methods design. In the quantitative phase, a small learner corpus of argumentative essays written by Indonesian EFL students was analysed to describe the distribution of grammatical and lexical cohesive devices and to examine their association with writing quality. In the qualitative phase, a subset of essays representing higher- and lower-rated texts underwent discourse analysis based on Halliday and Hasan's cohesion framework to explore how cohesion functioned in context. The research took place in the English Education Department of Universitas Ibrahimi Situbondo, Indonesia. From a population of 40 undergraduates enrolled in an academic writing course, 15 students were selected using simple random sampling. All participants had completed at least one semester of writing instruction and were familiar with basic argumentative structures. Institutional permission was obtained, students were informed that their work could be used for research without affecting their grades, and all texts were anonymised.

2.2 Data Collection and Coding

Data consisted of 15 individual argumentative essays produced as a regular course assignment on topics related to EFL/ESL learning. Students were instructed to state a clear position, support it with reasons and examples, and provide a conclusion. Essays were transcribed verbatim, lightly corrected only for obvious spelling errors, and the total number of words in each essay was calculated to allow frequency normalisation. Grammatical cohesion was coded for personal, demonstrative and comparative reference, articles, and conjunctions expressing additive, adversative, causal, and temporal relations. Lexical cohesion was coded for repetition, synonymy or near-synonymy, antonymy, and meronymy. All cohesive devices were identified manually and

recorded in coding sheets for each essay, and two indices of lexical cohesion were derived: the total number of lexical cohesive items and the number of different types used. Writing quality was rated using an analytic rubric covering focus, development, unity, coherence, and correctness on a three-point scale (good, enough, bad). These categories were converted to numerical values (3, 2, 1) and summed to yield an overall writing score (range 5–15). Two experienced writing instructors rated all essays; disagreements were discussed until consensus was reached, and the agreed scores were used for analysis.

2.3 Data Analysis

Quantitative analysis addressed two questions: the overall distribution of grammatical and lexical cohesion in the learner corpus and the extent to which cohesion indices related to writing quality. Descriptive statistics (means, standard deviations, minima, maxima) were calculated for all cohesion measures and writing scores. Because of the small sample size and non-normal distributions, Spearman's rank-order correlations were used to examine relationships between normalised cohesion frequencies (e.g. reference, conjunctions, total lexical cohesion, lexical diversity) and overall writing scores, with particular attention to coherence. For additional insight, essays were divided into higher- and lower-rated groups using a median split, and group differences in cohesion indices were inspected; effect sizes were considered when interpreting patterns. The qualitative phase involved an in-depth discourse analysis of four focal essays (two higher-rated and two lower-rated). Using Halliday and Hasan's cohesion framework, the analysis focused on reference chains, the signalling of logical relations through conjunctions, and the role of lexical reiteration in supporting or undermining thematic development and argument progression. Insights from this qualitative analysis were used to interpret and enrich the quantitative findings, thereby enhancing the trustworthiness and explanatory power of the study.

3. Result

3.1 Distribution of Grammatical and Lexical Cohesion

Before examining how cohesion relates to writing quality, it is important to obtain an overview of how often different cohesive devices occur in the learners' argumentative essays. This subsection therefore reports the distribution of grammatical and lexical cohesion across the 15 texts. Descriptive statistics were calculated for each cohesion category, including the mean frequency, standard deviation, and observed minimum and maximum values. These figures provide a baseline profile of the learners' typical cohesion patterns and reveal which cohesive resources they tend to rely on most.

Table 1. Descriptive statistics for grammatical cohesion indices.

Cohesion index	Mean	SD	Min	Max
Personal reference	6.07	3.75	2	13
Demonstrative reference	3.80	2.99	0	13
Comparative reference	0.73	0.68	0	2
Definite articles (a/an/the)	14.27	5.45	2	22
Conjunctions (all types)	10.53	3.83	7	20
Total grammatical devices	35.40	10.51	20	59

Table 1 presents the descriptive statistics for grammatical cohesion indices. Overall, the learners' essays contained an average of 35.40 grammatical cohesive devices per text, with considerable variation between essays (range 20–59 items). Among the individual categories, definite articles (a, an, the) were by far the most frequent, with a mean of 14.27 occurrences per essay. Conjunctions also appeared

frequently (M = 10.53), indicating that students regularly signalled some form of logical connection between clauses, although the qualitative analysis later shows that these connections were often limited to simple additive links. Personal reference (M = 6.07) and demonstrative reference (M = 3.80) occurred less often, and comparative reference was relatively rare (M = 0.73), suggesting that learners seldom expressed nuanced relations of similarity, difference, or degree. Taken together, these figures indicate a strong reliance on articles and basic conjunctions as the main grammatical resources for cohesion.

Table 2. Descriptive statistics for lexical cohesion indices (N = 15 essays).

Lexical cohesion index	Mean	SD	Min	Max
Total lexical cohesive devices	61.87	19.97	24	111
Lexical diversity index (no. of types*)	2.53	0.50	2	3

Table 2 summarises the distribution of lexical cohesion indices. On average, each essay contained 61.87 lexical cohesive devices, with a wide range from 24 to 111 items, which reflects substantial differences in how densely learners repeated or related lexical items in their texts. The lexical diversity index, which indicates how many different types of lexical cohesion (repetition, synonymy, antonymy, meronymy) were used in a given essay, had a mean value of 2.53 (SD = 0.50) and ranged from 2 to 3. This means that most essays relied on two or three kinds of lexical ties, typically dominated by simple repetition with occasional use of synonyms or antonyms, while meronymy appeared only sporadically. The combination of high overall lexical frequency and relatively low diversity suggests that learners tend to recycle the same words rather than exploiting a broader repertoire of semantically related expressions to develop their arguments.

3.2 Relationships Between Cohesion and Writing Quality

To examine whether learners who used more or more varied lexical cohesion also produced better essays, lexical indices were correlated with the analytic writing scores. In addition, essays were divided into higher- and lower-rated groups to compare their lexical cohesion profiles. This subsection reports the results of these analyses.

Table 3. Spearman correlations between lexical cohesion indices and writing quality

Cohesion index	ρ with overall writing score	p-value	ρ with coherence score	p-value
Total lexical cohesive devices	0.02	0.94	−0.43	0.11
Lexical diversity index*	−0.10	0.73	−0.38	0.16

*Lexical diversity index = number of lexical cohesion types (Repetition, Synonymy, Antonymy, Meronymy) used in each essay.

Table 3 presents the Spearman correlations between the lexical cohesion indices and the writing quality measures. The correlations with overall writing score were very small and non-significant: total lexical cohesive devices showed a near-zero relationship (ρ = 0.02, p = 0.94), and lexical diversity was also weakly and negatively related to overall quality (ρ = −0.10, p = 0.73). When focusing specifically on the coherence dimension, the coefficients were somewhat larger in magnitude but still non-significant. Interestingly, both total lexical devices (ρ = −0.43, p = 0.11) and lexical diversity (ρ = −0.38, p = 0.16) tended to correlate negatively with coherence, suggesting that essays with more frequent or more varied lexical ties were not necessarily perceived as more coherent. Taken together, these results indicate that the quantity of lexical cohesion alone does not straightforwardly translate into higher writing quality in this learner corpus.

Table 4. Lexical cohesion indices in higher- and lower-rated essays

Lexical cohesion index	Higher-rated essays (N = 5) M (SD)	Lower-rated essays (N = 4) M (SD)
Total lexical cohesive devices	56.20 (6.62)	53.50 (19.93)
Lexical diversity index[1]	2.40 (0.49)	2.50 (0.50)

[1]*Lexical diversity index = number of lexical cohesion types (repetition, synonymy, antonymy, meronymy) used in each essay.*

Table 4 compares the lexical cohesion indices for higher- and lower-rated essays. On average, higher-rated texts contained slightly more lexical cohesive devices (M = 56.20, SD = 6.62) than lower-rated texts (M = 53.50, SD = 19.93), but the difference in means was small and the lower-rated group showed much greater variability. A similar pattern emerged for lexical diversity: higher-rated essays used, on average, 2.40 types of lexical cohesion (SD = 0.49), while lower-rated essays used 2.50 types (SD = 0.50). Thus, the two groups did not differ meaningfully in either the density or the diversity of lexical cohesion. These findings reinforce the correlational results and suggest that what distinguishes stronger from weaker essays is not simply how many lexical ties are present, but how those ties are deployed functionally within the developing argument—a point that is explored further in the qualitative analysis.

3.3 Qualitative patterns in higher- and lower-rated essays

While the quantitative analyses showed that lexical cohesion indices were only weakly related to writing quality, a closer look at individual texts revealed clear qualitative contrasts between higher- and lower-rated essays. To explore how cohesion functioned in context, four focal essays were examined in detail: two representing the higher-rated group and two representing the lower-rated group. The comparison in Table 5 summarises the main differences in cohesion patterns across six aspects: reference chains, pronoun use, conjunctions and logical relations, argumentative flow, lexical reiteration, and thematic development.

Table 5. Qualitative comparison of cohesion patterns in higher- and lower-rated essays

Aspect of cohesion	Higher-rated essays (2 focal texts)	Lower-rated essays (2 focal texts)
Reference chains	Reference chains are generally sustained across clauses and paragraphs; key participants are tracked consistently.	Reference chains are often broken; shifts in reference occur without clear lexical or pronominal links.
Pronoun use and referential clarity	Personal pronouns (e.g. *they, it, this*) typically have clear antecedents; few ambiguous references.	Pronouns are sometimes used without explicit antecedents, creating ambiguity about who or what is being discussed.
Conjunction types and logical relations	A wider range of conjunctions is used to mark addition, contrast and cause–effect (e.g. *because, however, therefore*).	Logical relations are frequently signalled by the additive conjunction *and* alone; markers of contrast or justification are scarce.
Argumentative flow and coherence	Conjunctions and reference chains support a step-by-step progression from claim to reason to conclusion.	Ideas are loosely connected; missing or weak conjunctive links result in sudden topic shifts and fragmented arguments.

Lexical reiteration	Repetition of key terms is complemented by synonyms and related expressions, maintaining topic continuity while avoiding redundancy.	Lexical cohesion is dominated by unvaried repetition of the same nouns and adjectives, sometimes within the same sentence.
Thematic development	Reiterated and related lexical items help to build and extend themes across paragraphs, supporting a clear stance.	Repetition tends to recycle the same point rather than extending it, leading to circular or underdeveloped themes.

As summarised in Table 5, higher-rated essays tended to maintain more stable reference chains. Key participants such as "students", "teachers" or "online learning" were introduced and then consistently tracked across clauses and paragraphs through a mix of full noun phrases and pronouns. This continuity enabled readers to follow who or what was being talked about at each point in the text. In the lower-rated essays, by contrast, reference chains were frequently broken. New pronouns or noun phrases were introduced without clear links to earlier mentions, which sometimes forced the reader to infer connections or left the referent entirely ambiguous. This pattern suggests that referential cohesion in weaker essays was fragile, even when the writers employed a similar number of cohesive devices.

Related to this, pronoun use and referential clarity also distinguished the two groups. In the higher-rated essays, personal pronouns such as *they, it,* and *this* usually had explicit and easily recoverable antecedents, and shifts from one referent to another were signalled by lexical reformulation. In the lower-rated essays, pronouns were occasionally used without a clear antecedent or were positioned so far from their referents that their interpretation became uncertain. Such ambiguous pronominal reference undermined local coherence and contributed to the impression of "jumping" from one idea to another, even when the surface-level use of pronouns was frequent.

Differences in conjunction types and logical relations further illustrate how cohesion quality, rather than quantity, mattered. Higher-rated essays used a wider range of conjunctions to mark addition, contrast, and cause and effect relations, including items such as *because*, *however*, and *therefore*. These connectors guided the reader through sequences of reasons, counterarguments, and conclusions, making the argumentative structure more visible. In contrast, lower-rated texts relied heavily on the additive conjunction *and* as an all-purpose linker, with relatively few explicit markers of contrast or justification. As a result, relationships between propositions, such as whether a sentence provided a reason, an example, or a concession, often remained implicit and sometimes opaque.

These differences in reference and conjunction use were reflected in the overall argumentative flow and coherence. In the higher-rated essays, cohesive devices worked together to support a step-by-step progression from claim to reason to conclusion. Ideas were introduced, elaborated, and then tied back to the main stance, producing a sense of controlled development. In the lower-rated essays, ideas were more loosely connected. Missing or weak conjunctive links, combined with ruptured reference chains, led to sudden topic shifts and segments of text that read more like isolated statements than parts of a unified argument. This helps to explain why increased frequencies of lexical cohesion did not automatically yield higher coherence scores in the quantitative analysis.

The qualitative analysis also revealed notable contrasts in lexical reiteration. Higher-rated essays did rely on repetition of key terms, but this repetition was often complemented by synonyms and related expressions, which allowed writers to maintain

topic continuity while avoiding excessive redundancy. For example, a focal term such as "EFL students" might alternate with "learners", "they", and "students in Indonesia", creating a cohesive lexical field around the central theme. In lower-rated essays, lexical cohesion was dominated by unvaried repetition of the same nouns and adjectives, sometimes several times within a single sentence. This mechanical recycling of vocabulary did little to develop ideas and occasionally produced a monotonous textual effect.

4. Discussion

This study examined how Indonesian EFL undergraduates use grammatical and lexical cohesion in argumentative essays and how these patterns relate to analytic ratings of writing quality. Overall, the findings suggest a nuanced picture: learners made extensive use of cohesive devices, especially articles, conjunctions, and lexical repetition, yet simple frequency and diversity indices of lexical cohesion showed only weak, and sometimes negative, associations with overall scores and coherence. In contrast, the qualitative analysis revealed clear differences between higher- and lower-rated essays in the functional deployment of cohesion, particularly in reference chains, conjunctive relations, and lexical reiteration. Together, these results support the view that cohesion is important for writing quality, but that raw counts of cohesive ties are insufficient to explain raters' judgments (Faigley & Witte, 1981; McNamara, Crossley, & McCarthy, 2010).

The distributional patterns are broadly consistent with previous research on cohesion in L2 writing. Learners relied heavily on grammatical devices such as articles and basic conjunctions and on lexical repetition, while more nuanced resources, such as comparative reference and varied lexical relations, were underused. Similar tendencies have been reported in other learner corpora, where L2 writers initially depend on a restricted repertoire of cohesive devices and only gradually diversify their cohesive resources (Neary-Sundquist, 2013). Longitudinal studies further show that global cohesion develops slowly and non-linearly, with considerable individual variation in learners' trajectories (Zhang & Zhang, 2024). Our corpus, although small and cross-sectional, aligns with this developmental picture: students can produce dense cohesive chains, but those chains are often built from a narrow set of forms.

The weak and non-significant correlations between lexical cohesion indices and writing quality mirror long-standing debates about the relationship between cohesion and coherence. Classic critiques argue that cohesion counts should not be treated as direct proxies for coherence or quality, because readers' evaluations also depend on content, organisation, and discourse-level logic. More recent work likewise reports that cohesion indices often have limited power to discriminate high- and low-quality essays (McNamara et al., 2010; Spiegel & Fitzgerald, 1990) and that raters tend to privilege overall coherence and rhetorical effectiveness over the sheer number of cohesive ties (Johnson, 1992). Our results fit this pattern: essays with more frequent or more "diverse" lexical devices were not necessarily judged as more coherent. In fact, the slightly negative correlations with coherence suggest that dense repetition, when not managed strategically, may make texts feel redundant or circular rather than well structured.

The qualitative analysis helps explain why cohesion frequency alone was a poor predictor of quality. Higher-rated essays tended to maintain stable reference chains, with pronouns and noun phrases clearly tracking key participants and concepts across clauses and paragraphs. They also used a broader range of conjunctions—such as causal and adversative linkers—to signal argumentative relations, and combined repetition with synonyms and related expressions to build thematic chains. These

patterns resonate with studies showing that effective academic writing relies on lexical cohesion to develop topics and connect propositions (Stotsky, 1983; Mojica, 2008; Kuo, 1995), and that high-rated texts typically display more controlled and purposeful use of cohesive devices (Bacha, 2002). By contrast, lower-rated essays in our corpus were marked by broken reference chains, over-reliance on the additive conjunction *and*, and unvaried repetition that recycled points rather than extending them. In this sense, our findings echo broader observations from the Indonesian EFL context that students often draw on a limited range of cohesive expressions and structuring signals when constructing arguments in English (Arsyad, 1999; Oktavianti & Prayogi, 2022; Rusfandi, 2015).

These results have clear pedagogical implications. If the key issue is not simply "more cohesion" but more strategic cohesion, instruction needs to move beyond asking students to insert linkers or avoid repetition in a mechanical way. Genre-based and explicit pedagogies can help by foregrounding how cohesive devices support the communicative purposes and stages of argumentative writing (Hyland, 1990; Bejarano & Chapetón, 2013). In practical terms, teachers might engage students in analysing model essays to trace reference chains, identify how different conjunctions encode specific logical relations, and map lexical networks around key themes. Teaching activities that require learners to repair ambiguous pronouns, choose more appropriate conjunctions, or paraphrase key terms could foster a richer and more functional command of cohesion. At the same time, raising awareness of cohesion as part of broader argumentative competence aligns with calls to integrate writing instruction with the development of critical thinking and disciplinary reasoning (Ferretti & Graham, 2019; Zaheer, Rahman, & Sharma, 2024).

Several limitations should be acknowledged. The corpus was small and drawn from a single institution, limiting the generalisability of the quantitative patterns. The analysis focused on selected cohesion categories and did not systematically assess accuracy or appropriateness, which may also influence raters' perceptions. Moreover, the cross-sectional design cannot capture the developmental dynamics highlighted in longitudinal work (Zhang & Zhang, 2024; Tabari & Wind, 2023). Future studies could address these limitations by including larger and more diverse samples, combining cohesive indices with measures of syntactic complexity and lexis, and tracking changes in cohesion alongside instructional interventions. Despite these constraints, the present study contributes to ongoing debates by showing that, in this Indonesian EFL context, stronger essays are distinguished less by the amount of cohesion they contain than by how cohesively they guide readers through an argument.

5. Conclusion

This study set out to examine how Indonesian EFL undergraduates use grammatical and lexical cohesion in argumentative essays and how these cohesion patterns relate to analytic ratings of writing quality. The quantitative analysis showed that the learners produced a relatively high density of cohesive devices, particularly through articles, conjunctions, and lexical repetition. However, indices of lexical cohesion, both in terms of total frequency and diversity, displayed only weak and non-significant relationships with overall writing scores and coherence. In other words, having "more" lexical cohesion did not straightforwardly translate into better-rated essays.

The qualitative analysis helped to clarify this apparent paradox. Higher-rated essays were not distinguished by the sheer amount of cohesion, but by the strategic and functional deployment of cohesive devices. These texts typically sustained clear reference chains, used a wider range of conjunctions to signal argumentative relations,

and combined repetition with synonyms and related expressions to build and extend themes. Lower-rated essays, by contrast, often showed broken reference chains, heavy reliance on the additive conjunction and, and unvaried lexical repetition that recycled ideas rather than developing them. These contrasts suggest that it is the quality, not the quantity, of cohesion that plays a more decisive role in raters' judgments of argumentative writing.

Pedagogically, the findings point to the need to move beyond treating cohesion as a checklist of linguistic forms to be added to texts. Instruction should instead foreground cohesion as a meaning-making resource that supports argument structure, thematic progression, and reader guidance. Activities that help students trace and repair reference chains, choose conjunctions that accurately encode logical relations, and build lexical networks around key claims may be particularly beneficial. At the same time, the study's small corpus and single-institution context limit the generalisability of the findings. Future research with larger and more diverse learner populations, and with longitudinal or intervention designs, is needed to further unpack how cohesion develops over time and how targeted teaching can foster more effective cohesive control in EFL argumentative writing.

References

Abdollahzadeh, E., Abdollahzadeh, E., Farsani, M. A., & Beikmohammadi, M. (2017). Argumentative writing behavior of graduate EFL learners. Argumentation. https://doi.org/10.1007/S10503-016-9415-5

Arsyad, S. (1999). The Indonesian and English argument structure: A cross-cultural rhetoric of argumentative texts. Australian Review of Applied Linguistics. https://doi.org/10.1075/ARAL.22.2.06ARS

Bacha, N. N. (2002). Academic lexical literacy: Investigating the cohesion of Arabic speakers' essays in English. International Journal of Arabic-English Studies. https://doi.org/10.33806/ijaes2000.3.1.9

Barrot, J., & Gabinete, M. K. L. (2021). Complexity, accuracy, and fluency in the argumentative writing of ESL and EFL learners. International Review of Applied Linguistics in Language Teaching. https://doi.org/10.1515/IRAL-2017-0012

Bejarano, P. A. C., & Chapetón, C. M. (2013). The role of genre-based activities in the writing of argumentative essays in EFL. Profile Issues in Teachers' Professional Development.

Colomb, G. G., & Griffin, J. A. (2004). Coherence on and off the page: What writers can know about writing coherently. New Literary History. https://doi.org/10.1353/NLH.2004.0027

Crowhurst, M. (1987). Cohesion in argument and narration at three grade levels. Research in the Teaching of English.

Dornbrack, J., & Dixon, K. (2014). Towards a more explicit writing pedagogy: The complexity of teaching argumentative writing. Reading and Writing. https://doi.org/10.4102/rw.v5i1.40

Faigley, L., & Witte, S. P. (1981). Coherence, cohesion, and writing quality. College Composition and Communication. https://doi.org/10.2307/356693

Ferretti, R. P., & Graham, S. (2019). Argumentative writing: Theory, assessment, and instruction. Reading and Writing. https://doi.org/10.1007/S11145-019-09950-X

Huang, D., & Watzinger–Tharp, J. (2023). The use of cohesive devices as proficiency level discriminators in Chinese DLI learners' writing. Foreign Language Annals. https://doi.org/10.1111/flan.12722

Hyland, K. (1990). A genre description of the argumentative essay. RELC Journal. https://doi.org/10.1177/003368829002100105

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

Johnson, P. (1992). Cohesion and coherence in compositions in Malay and English. RELC Journal. https://doi.org/10.1177/003368829202300201

Kro, C.-H. (1995). Cohesion and coherence in academic writing: From lexical choice to organization. RELC Journal. https://doi.org/10.1177/003368829502600103

Kuhn, D. (2019). Critical thinking as discourse. Human Development. https://doi.org/10.1159/000500171

Mahfoudhi, A. (2003). Writing processes of EFL students in argumentative essays: A case study. ITL – International Journal of Applied Linguistics. https://doi.org/10.2143/ITL.139.0.2003202

Mallahi, O. (2024). Exploring the status of argumentative essay writing strategies and problems of Iranian EFL learners. Asian-Pacific Journal of Second and Foreign Language Education. https://doi.org/10.1186/s40862-023-00241-1

McCulley, G. A. (1985). Writing quality, coherence, and cohesion. Research in the Teaching of English.

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. Written Communication. https://doi.org/10.1177/0741088309351547

Mojica, L. A. (2008). Reiterations in ESL learners' academic papers: Do they contribute to lexical cohesiveness? Asia-Pacific Education Researcher. https://doi.org/10.3860/TAPER.V15I2.83

Neary-Sundquist, C. A. (2013). The development of cohesion in a learner corpus. Studies in Second Language Learning and Teaching. https://doi.org/10.14746/SSLLT.2013.3.1.6

Oktavianti, I. N., & Prayogi, I. (2022). Discourse functions of lexical bundles in Indonesian EFL learners' argumentative essays: A corpus study. Studies in English Language and Education. https://doi.org/10.24815/siele.v9i2.23995

Palmer, J. C. (1999). Coherence and cohesion in the English language classroom: The use of lexical reiteration and pronominalisation. RELC Journal. https://doi.org/10.1177/003368829903000204

Putra, J. W. G., Teufel, S., & Tokunaga, T. (2023). Improving logical flow in English-as-a-foreign-language learner essays by reordering sentences. Artificial Intelligence. https://doi.org/10.1016/j.artint.2023.103935

Rhetorical pattern of the Indonesian EFL undergraduate students' writings. (2022). Studies in English Language and Education. https://doi.org/10.24815/siele.v9i1.13640

Rusfandi, R. (2015). Argument–counterargument structure in Indonesian EFL learners' English argumentative essays: A dialogic concept of writing. RELC Journal. https://doi.org/10.1177/0033688215587607

Sandra, R. P., Hwang, W.-Y., Zafirah, A., Hariyanti, U., Engkizar, E., Hadi, A., & Fauzan, A. (2024). Crafting compelling argumentative writing for undergraduates: Exploring the nexus of digital annotations, conversational agents, and collaborative concept maps. Journal of Educational Computing Research. https://doi.org/10.1177/07356331241242437

Spiegel, D. L., & Fitzgerald, J. (1990). Textual cohesion and coherence in children's writing revisited. Research in the Teaching of English.

Stotsky, S. (1983). Types of lexical cohesion in expository writing: Implications for developing the vocabulary of academic discourse. College Composition and Communication. https://doi.org/10.2307/357899

Tabari, M. A., & Wind, A. M. (2023). Dynamic development of cohesive devices in English as a second language writing. Iral-International Review of Applied Linguistics in Language Teaching. https://doi.org/10.1515/iral-2022-0205

Tierney, R. J., & Mosenthal, J. H. (1983). Cohesion and textual coherence. Research in the Teaching of English.

Venegas, R., Castro-Cano, E. M., & Cornejo, D. L. (2024). Marcadores discursivos y calidad de la escritura: Estudio empírico-computacional del ensayo en estudiantes universitarios. Ikala, Revista de Lenguaje y Cultura. https://doi.org/10.17533/udea.ikala.355065

Yasuda, S. (2023). What does it mean to construct an argument in academic writing? A synthesis of English for general academic purposes and English for specific academic perspectives. Journal of English for Academic Purposes. https://doi.org/10.1016/j.jeap.2023.101307

Zaheer, S., Rahman, M. M., & Sharma, A. (2024). Awareness of discourse features in teaching writing skills. Jurnal Ilmu Kemanusiaan. https://doi.org/10.21315/kajh2024.31.1.1

Zanardi, N. (1994). Cohesion in Italian adult learners' and native speakers' compositions: A comparison. Australian Review of Applied Linguistics. https://doi.org/10.1075/ARAL.17.2.02ZAN

Zhang, J. (2023). Lexical cohesion development in English as a foreign language learners' argumentative writing: A latent class growth model approach. Linguistics and Education. https://doi.org/10.1016/j.linged.2023.101255

Zhang, J., & Zhang, L. (2024). Examining L2 students' development of global cohesion and its relationship with their argumentative essay quality. Iral-International Review of Applied Linguistics in Language Teaching. https://doi.org/10.1515/iral-2023-0069