

Beyond Linguistic Similarity: Conceptual-Linguistic Divergence in EFL Authorship Analysis

Amirudin Latif^{1*}, Fitri Palupi Kusumawati², Eva Faliyanti³

Universitas Muhammadiyah Metro, Indonesia

amirudinlatif80@gmail.com

fitripalupi.1985@ummetro.ac.id

faliyantieva@gmail.com

ABSTRACT

Authorship consistency has taken center stage in academic writing today, especially with the rise of AI and digital writing tools. While traditional authorship studies mostly look at linguistic similarities between disputed and genuine texts, the role of conceptual continuity remains largely overlooked in educational settings. This study addresses this gap by examining authorship consistency in EFL academic writing. Using a qualitative forensic linguistic case study, we analyzed one undergraduate thesis abstract (the questioned document) alongside eight known texts written by the same student. The framework combined linguistic profiling—looking at lexical, grammatical, syntactic, and stylistic features—with a conceptual analysis driven by a handwritten abstract reproduction task. Interestingly, the findings revealed a sharp linguistic mismatch between the thesis abstract and the student's usual writing profile. However, the student displayed deep conceptual continuity, flawlessly rebuilding the study's topic, goals, methodology, findings, and conclusions. This striking contrast where language diverges but the core concept aligns is what we define as *conceptual-linguistic divergence*. Ultimately, this study suggests that educational authorship checks can no longer rely solely on language patterns; they must integrate conceptual evidence, particularly in EFL contexts.

Keywords: authorship consistency; forensic linguistics; EFL academic writing; conceptual continuity; conceptual-linguistic divergence

INTRODUCTION

Academic writing is the cornerstone of higher education, acting as the ultimate proof of a student's critical thinking, disciplinary knowledge, and academic capability. It requires students to do heavy lifting: evaluating scholarly sources, forming cohesive arguments, and adopting specific academic conventions. This benchmark becomes exceptionally steep in EFL (English as a Foreign Language) settings. Here, students face the demanding task of balancing rich content and complex organization, all while trying to maintain linguistic accuracy in a language they are still mastering. (Hyland, 2003, 2019). This demonstrates that academic writing is not just about language proficiency, but also deep intellectual engagement, which is why it remains a cornerstone of educational assessment. Traditionally, schools and universities have operated on a simple premise: a student's paper is a direct, authentic reflection of their own knowledge and language skills. However, this foundational assumption is cracking under the weight of the modern educational landscape. Today, students rarely write in a vacuum; they naturally collaborate with supervisors and peers, and lean on editors,

digital dictionaries, translation software, and advanced writing assistants. Consequently, the lines around authorship, authenticity, and intellectual ownership have blurred, turning these concepts into some of the most pressing challenges in higher education today.

The rise of generative artificial intelligence (AI) has pushed the evolution of academic writing even further. Today, it is common for students to lean on AI tools to polish their grammar, enrich their vocabulary, organize their thoughts, summarize dense data, or even draft whole blocks of text. While these applications provide excellent scaffolding for learning and honing writing skills, they simultaneously disrupt conventional boundaries. As a result, they force educators to rethink what terms like "authorship" and "authenticity" truly mean in a digitized academic world. (Kasneci et al., 2023; Lo, 2023). Recent academic debates reflect an urgent skepticism about modern student writing. There is a growing anxiety over whether texts heavily shaped by AI can still be considered a true measure of a student's solo capabilities and unique intellectual footprint. (Cotton et al., 2024; Perkins, 2023). Consequently, the focus of academic honesty has shifted dramatically. It is no longer just about running papers through plagiarism software; the debate has matured into a deeper investigation into the actual relationship between the writer and the text, and the true authenticity of ownership. (Moorhouse et al., 2023; Tlili et al., 2023).

To determine the true identity behind a piece of writing, forensic linguistics provides highly effective analytical methods. Because it analyzes language as a form of evidence, it has routinely been used to handle everything from copy-pasted work and nameless letters to legal document verification and copyright battles. (Coulthard & Johnson, 2010; Fobbe, 2011; McMenamin, 2002). Forensic authorship analysis operates on the premise that writers maintain distinct and consistent linguistic habits. These patterns surface through predictable choices in vocabulary, grammar, syntax, and overall style. By systematically evaluating these specific features, researchers can map out similarities and differences between texts to determine potential authorship.

Over the last twenty years, the field of authorship attribution has come a long way. Early research in this area mostly focused on things like linguistic profiling, style analysis, and stylometrics to figure out who actually wrote a text. (Juola, 2006; Stamatatos, 2009). Newer studies build on these ideas by introducing the concept of linguistic individuality. The main argument here is that everyone has their own distinct, repeating writing habits, which can act almost like a fingerprint to identify the author. (Nini, 2023). Despite recent methodological breakthroughs, the core of most authorship investigations remains unchanged. Analysts still primarily evaluate writer-text relationships by comparing linguistic similarities between the disputed document and known reference samples.

Although linguistic profiling remains a cornerstone of authorship analysis, its implementation within educational environments introduces distinct difficulties. Studies in second-language acquisition consistently reveal that a learner's conceptual understanding and their grammatical or linguistic proficiency do not always advance at the same pace. (Hyland & Jiang, 2021). Students might fully understand complex course concepts but still lack the advanced writing skills

needed to express that knowledge in a formal academic tone. Flowerdew (2020) echoes this sentiment, arguing that the transition from abstract thought to formal academic writing is often difficult for multilingual individuals. Therefore, relying strictly on language proficiency as a metric can overlook the writer's actual intellectual contribution and grasp of the subject.

This mismatch creates a real dilemma for standard authorship verification in EFL contexts. Imagine a student who struggles with basic English on a timed test, but then submits a beautifully written term paper. Standard detection methods would automatically conclude the student didn't write it. Yet, a jump in linguistic quality does not prove the ideas are not theirs. Someone can completely own, understand, and explain the intellectual heavy lifting behind a text, even if they lack the vocabulary to recreate its polished style on demand.

Current authorship research suffers from a major gap: it rarely explores the link between conceptual continuity and linguistic realization in school settings. Instead, existing literature relies almost entirely on tracking linguistic similarities to identify an author. This leaves the writer's familiarity with the actual content largely unexamined. Such an omission is highly problematic for EFL environments, where language performance and intellectual comprehension rarely develop at the exact same pace.

Table 1. Research Gap Analysis

Recent Studies	Main Focus	Limitation
Kasneci et al. (2023)	AI opportunities and challenges in education	Limited discussion of forensic authorship evidence
Perkins (2023)	Academic integrity and generative AI	Focuses on integrity issues rather than authorship consistency
Lo (2023)	ChatGPT in education	Limited examination of writer–text relationships
Cotton et al. (2024)	Academic integrity in AI-assisted writing	Emphasizes policy and assessment implications
Nini (2023)	Linguistic individuality and authorship analysis	Focuses primarily on linguistic evidence
Present Study	Linguistic and conceptual evidence in EFL authorship investigation	Introduces conceptual-linguistic divergence

The research trends highlighted in Table 1 prove that scholars are deeply focused on linguistic authorship analysis, AI writing tools, and academic integrity. Yet, almost all of these studies still rely strictly on tracking grammar and vocabulary matches. This creates a massive blind spot when a student clearly knows the subject matter inside out, but writes in a style that doesn't match the paper in question. We still know very little about how this gap between conceptual grasp and language differences plays out, especially for EFL students.

This study addresses the unexamined area by tracking authorship consistency in EFL writing through a forensic lens, cross-examining one disputed document with eight reference samples from the same individual. Crucially, the

methodology doesn't stop at linguistic profiling. By incorporating a task where the participant reproduces the abstract by hand, this project specifically isolates and measures conceptual continuity to see if the ideas truly belong to the writer.

The evidential patterns uncovered in this data lead us to introduce a new conceptual framework: conceptual-linguistic divergence. This term characterizes scenarios where a profound continuity of ideas exists alongside a sharp mismatch in language style. By doing so, this study challenges the conventional reliance on surface-level text similarity during authorship disputes. Instead, it offers educational investigators a much broader, more sophisticated toolkit for evaluating writer-text relationships.

Accordingly, the study seeks to answer the following research questions:

RQ1. What linguistic features characterize the participant's known documents?

RQ2. What linguistic features characterize the questioned document?

RQ3. To what extent does the questioned document demonstrate linguistic consistency with the known documents?

RQ4. How does conceptual-linguistic divergence contribute to the interpretation of authorship consistency in EFL academic writing?

METHODS

Research Design

The methodology of this research relies on a qualitative forensic linguistic case study design to evaluate consistency in EFL writing. Utilizing a case-study framework is justified given the project's focus: a granular, comprehensive analysis of a single authorship dispute featuring one questioned text and a collection of baseline documents from the same subject. The study prioritizes a nuanced interpretation of the interplay between conceptual and linguistic evidence over statistical generalization within educational contexts.

In tracking the author's identity, the study mirrors the standard questioned-document procedures typically utilized in forensic linguistics. (McMenamin, 2002; Coulthard & Johnson, 2010). Beyond conventional linguistic comparison, the methodological framework incorporated a handwritten abstract reproduction task to isolate and evaluate conceptual continuity. This dual-layered research design provided an empirical basis to scrutinize both the surface-level linguistic realization and the underlying conceptual comprehension demonstrated by the participant.

Participant and Data Sources

The subject under investigation was an undergraduate student majoring in English Education at an Indonesian tertiary institution. A purposive sampling technique was employed to select this participant, primarily driven by the availability of an extensive writing corpus necessary to construct a reliable baseline linguistic profile and execute the subsequent textual comparison.

A total of nine documents were gathered to form the study's dataset, consisting of one questioned document (QD) and eight known documents (KDs). While the QD was derived from the participant's undergraduate thesis abstract, seven of the comparative texts (KD2–KD8) were compiled from prior independently written materials, such as translation activities, classroom essays, grammar

worksheets, and academic prose. To supplement these baseline texts, an additional sample (KD1) was generated via an experimental handwritten abstract reproduction task.

Table 2. Data Sources

Code	Document Type	Function in Analysis
QD	Undergraduate thesis abstract	Questioned document
KD1	Handwritten abstract reproduction task	Conceptual evidence
KD2–KD8	Independent writing samples	Linguistic evidence

Gathering multiple known documents allowed the investigation to establish a more authentic, long-term mapping of the writer's idiolect. Consequently, this minimized the risk of task-related skewing and ensured that the subsequent stylistic analysis reflected habitual rather than isolated writing behaviour.

Data Collection Procedures

The data gathering process was executed in a two-tier sequence. The initial phase involved compiling and systematically cataloging both the questioned text and the reference documents into a unified comparative corpus. These baseline materials were intentionally selected because they captured the subject's authentic, unassisted prose, thereby reflecting genuine language habits across diverse academic assignments. To maintain strict research ethics, complete anonymization was applied to all documents, ensuring the absolute confidentiality of the participant.

In the second stage of data collection, KD1 was elicited through a specialized handwritten abstract reproduction task. The participant was required to recall and rewrite the underlying arguments of the undergraduate thesis abstract from memory, without the aid of the original manuscript or secondary materials. No strict temporal constraints were enforced; instead, the participant was directed to focus primarily on capturing the intellectual content rather than duplicating the exact lexicon of the initial document.

The reproduction task was implemented not to measure memory retention, but rather to function as an elicitation tool for assessing familiarity with the conceptual framework of the questioned document. Specifically, the task isolated the participant's ability to recall and structurally reconstruct the essential components of the abstract, which included the research topic, thesis objective, methodology, key findings, and final conclusions. Prior to data collection, informed consent was obtained from the participant. Participation was voluntary, and all identifying information was removed during the analysis and reporting stages.

Analytical Framework

The study integrated two complementary forms of evidence: linguistic evidence and conceptual evidence. Linguistic evidence was examined through four dimensions commonly employed in forensic linguistic profiling: lexical features, grammatical features, syntactic features, and stylistic features. These dimensions were used to establish the participant's linguistic profile based on KD2–KD8 and

to evaluate the degree of correspondence between the known documents and the questioned document.

Conceptual evidence was examined through analysis of the handwritten reproduction task (KD1). Particular attention was given to the participant's ability to reconstruct the principal components of the thesis abstract, including the research topic, objective, methodology, findings, and conclusion.

Table 3. Analytical Indicators

Construct	Indicator
Lexical Consistency	Similar lexical patterns across documents
Grammatical Consistency	Similar grammatical tendencies across documents
Syntactic Consistency	Similar sentence-construction patterns
Stylistic Consistency	Similar discourse organization and writing style
Conceptual Continuity	Reconstruction of topic, objective, methodology, findings, and conclusion
Conceptual-Linguistic Divergence	Strong conceptual continuity accompanied by substantial linguistic divergence

The analytical model underlying the study is presented in Figure 1.

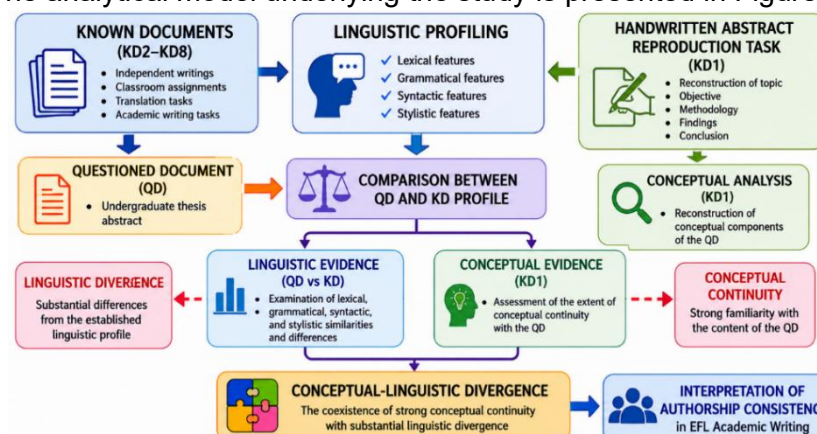


Figure 1 delineates how this study merges linguistic profiling and authorship consistency with conceptual continuity under a singular forensic framework. Operationally, the analysis commences by tracking the known documents (KD2–KD8) to isolate the participant's stable writing habits, drawing from recurring patterns in vocabulary choices, grammar preferences, syntax, and style. Once compiled, this profile functions as the primary evidential baseline against which authorship consistency is tested.

Following the baseline construction, the questioned document (QD) is cross-examined against the verified linguistic profile to gauge the degree of textual alignment between both datasets. This comparative process yields the primary linguistic evidence necessary to diagnose whether the language patterns remain consistent or exhibit substantial divergence.

Simultaneously, the handwritten abstract reproduction task (KD1) is analyzed to measure conceptual continuity. Sidestepping surface-level linguistic realization, this analytical phase focuses strictly on the participant's capacity to reconstruct the core architecture of the questioned document, specifically isolating

key components such as the research framework, primary objective, methodology, empirical findings, and final conclusion.

In the final phase, linguistic and conceptual data streams are integrated to formulate a holistic interpretation of the writer–text relationship. By examining the interaction between these two forms of evidence, the researcher can systematically identify patterns of conceptual-linguistic divergence. This analytical category identifies instances where substantial thematic and conceptual continuity is maintained despite significant stylistic and linguistic variance.

Data Analysis Procedures

A five-stage data analysis procedure was implemented to evaluate the compiled corpus. First, the baseline known documents (KD2–KD8) were analyzed to map out the participant's idiosyncratic linguistic profile. Second, the same analytical metrics were applied to evaluate the questioned document. Third, a comparative analysis was performed to measure the linguistic correspondence between the questioned text and the baseline profile. Fourth, the handwritten abstract reproduction task was evaluated to assess structural and conceptual continuity. The process culminated in the integration of both linguistic and conceptual datasets to formulate a holistic interpretation of the writer–text relationship.

Trustworthiness

Several measures were implemented to secure the trustworthiness and credibility of this investigation. First, the construction of the subject's linguistic profile utilized multiple baseline texts instead of a singular writing sample, thereby neutralizing task-specific biases. Second, comparative uniformity was achieved by applying the same analytical metrics across all compiled documents. Third, descriptive transparency was prioritized by embedded representative textual clips within the results, enabling external readers to cross-examine the evidential foundation supporting the qualitative conclusions.

Methodological triangulation was further operationalized by intersecting linguistic profiling with conceptual analysis. This integration of two complementary evidentiary data streams enhanced the interpretive credibility of the outcomes, facilitating a more nuanced and comprehensive evaluation of the participant's interaction with the questioned document.

FINDINGS AND DISCUSSION

Establishing a Stable Linguistic Baseline

The study first set out to see if the participant's known documents (KD2–KD8) showed enough linguistic consistency to build a reliable authorship profile. Interestingly, the analysis uncovered recurring patterns in vocabulary, grammar, sentence structure, and style, even when the writing tasks and purposes differed.

Lexically, the participant stuck to common words and straightforward language structures. They heavily repeated basic terms like *important*, *good result*, *many students*, and *learning English* throughout their writing. Advanced academic wording was scarce, usually replaced by simpler alternatives. This stylistic choice

suggests the author's primary goal was basic clarity rather than sophisticated vocabulary.

In terms of grammar, the findings showed repeating errors in subject–verb agreement, article usage, verb forms, and noun structures. For example:

“The students is interested to learn English because it give many benefit.”

These same grammatical features were found in multiple reference texts, suggesting they reflect the writer's ongoing habits rather than isolated performance errors. Such consistency matters because forensic linguistics looks for established patterns of language behaviour rather than occasional oversights. (McMenamin, 2002; Nini, 2023). At the syntactic level, the participant showed a clear preference for straightforward sentence forms, with very little use of subordination. Ideas were typically chained together through independent clauses and coordinating conjunctions. Stylistically, the texts aimed for direct communication and straightforward information sharing over stylistic flair or formal disciplinary jargon.

Taken as a whole, these repeating characteristics point to the existence of a steady linguistic profile. The steady consistency across these various texts reinforces the assumption that writers form unique language habits that act as markers for authorship identification. Consequently, KD2–KD8 served as a reliable benchmark against which the disputed text could be tested.

A Questioned Document That Does Not Fit the Established Profile

Analysis of the questioned document (QD) showed a linguistic profile that was vastly different from the participant's known baseline. Across word choice, grammar, sentence structure, and style, the disputed text displayed a much higher level of language sophistication.

For example:

“The findings revealed that the implementation of communicative language teaching significantly enhanced students' speaking performance.”

Terms like *implementation*, *significantly enhanced*, and *speaking performance* were completely missing from the participant's known writing. Similar trends were found throughout the entire abstract, showing a much more advanced academic tone and sharper precision in word choice.

Grammatically, the questioned document was highly accurate, and the errors that frequently appeared in the known texts were mostly missing. On top of that, the abstract utilized more intricate sentence frameworks, incorporating subordinate clauses, nominalizations, and subject-specific academic phrasing.

From a traditional forensic linguistic standpoint, these disparities point to a weak match between the disputed text and the writing the participant produced on their own. If authorship were judged strictly on language similarities, the questioned document would stand out as completely incompatible with the participant's known linguistic profile.

However, a mere shift in language style is enough to fully clarify the participant's connection to the document. Academic writing frequently involves drafts, edits, feedback, and outside language assistance—especially for those learning English as a foreign language. Because of this, we need more evidence before drawing any firm conclusions about authorship.

Evidence of Linguistic Divergence

A straight comparison between the disputed text and the participant's language baseline showed clear, consistent differences across every single area of analysis.

Table 4. Summary of Linguistic Findings

Dimension	Known Documents (KD2–KD8)	Questioned Document (QD)
Lexical Features	High-frequency vocabulary	Academic vocabulary
Grammatical Features	Recurring grammatical errors	High grammatical accuracy
Syntactic Features	Predominantly simple structures	Frequent complex constructions
Stylistic Features	Direct communication style	Formal academic discourse

The observed divergence was not confined to sporadic linguistic features but remained evident throughout the whole text. Rather, the questioned document was distinctly separated from the participant's established baseline by its sophisticated vocabulary, grammatical precision, syntactic complexity, and overall rhetorical framework.

These results match what earlier studies found. Basically, when writing styles are too different, it's a lot harder to confidently guess who wrote a text just by looking at the words. (Juola, 2006; Stamatatos, 2009). However, this study departs from conventional authorship research by treating textual evidence as merely a single component. It goes a step further by evaluating conceptual evidence derived from the handwritten copying task.

Emergence of Conceptual Continuity

The conceptual evidence revealed a sharp contrast in the findings. Even though the participant failed to match the document's advanced linguistic style, the handwritten reproduction proved they deeply understood the intellectual content of the abstract.

Table 5. Summary of Conceptual Continuity

Component	Successfully Reconstructed
Research Topic	Yes
Research Objective	Yes
Methodology	Yes
Findings	Yes
Conclusion	Yes

For instance, the original abstract stated:

"The purpose of this study was to investigate the effect of communicative language teaching on students' speaking performance."

The participant reproduced the central idea as:

"This research discussed communicative teaching and students speaking ability."

Even though the rewritten version used simpler wording and sentence structures, the core conceptual meaning remained intact. The participant

maintained this accurate representation across all major sections, including the methodology, findings, and conclusion.

Reconstructing these components shows that the participant thoroughly understood the document, rather than just knowing it superficially. They proved they could grasp the study's complex structure and ideas, supporting L2 writing research which shows that conceptual and linguistic skills don't always develop at the same rate. (Hyland, 2019; Hyland & Jiang, 2021; Flowerdew, 2020). The study therefore disputes the assumption that text style alone is enough to prove true authorship or intellectual ownership.

Conceptual-Linguistic Divergence

Combining linguistic and conceptual evidence brought the study's main conclusion to light. Linguistically, the questioned document was vastly different from the participant's usual writing profile. Conceptually, however, there was strong alignment, as the participant successfully reproduced the abstract's intellectual content. This intersection of robust conceptual alignment and major linguistic mismatch is defined here as conceptual-linguistic divergence. This term describes scenarios where the participant proves she fully understands and owns the core ideas, yet uses language that looks completely different from the style found in the questioned text.

Importantly, the findings suggest that authorship consistency cannot always be interpreted solely through linguistic similarity. In EFL academic writing, conceptual understanding and linguistic realization may develop unevenly. As a result, writers may retain ownership of ideas while lacking the linguistic resources necessary to reproduce the same level of textual sophistication independently. Unlike conventional authorship studies that primarily interpret writer-text relationships through linguistic correspondence, the present study demonstrates the value of integrating conceptual evidence into authorship investigations. The findings suggest that conceptual continuity may function as a complementary form of evidence when linguistic evidence alone produces inconclusive or potentially misleading interpretations.

Implications for Educational Authorship Investigation

The findings have important implications for educational authorship investigations. First, they demonstrate that linguistic divergence should not automatically be interpreted as evidence against authorship consistency. Educational writing often involves varying levels of assistance, revision, and language support, particularly in EFL contexts.

Second, the findings highlight the importance of examining conceptual evidence alongside linguistic evidence. The participant's ability to reconstruct the content of the questioned document suggests that intellectual ownership may remain detectable even when linguistic realization differs substantially.

Consequently, educational institutions may benefit from adopting multidimensional approaches to authorship evaluation that incorporate both linguistic and conceptual evidence.

Implications for Academic Integrity in the AI Era

The growing use of AI-assisted writing technologies further increases the importance of multidimensional authorship assessment. Contemporary writing tools can modify grammar, vocabulary, and discourse organization without necessarily altering a writer's conceptual understanding. As a result, textual sophistication alone may no longer provide a reliable indicator of intellectual ownership.

The present findings suggest that future academic-integrity frameworks should move beyond exclusive reliance on textual similarity and consider how conceptual evidence can complement linguistic evidence in evaluating writer-text relationships. Such an approach may provide a more balanced and educationally appropriate understanding of authorship in increasingly technology-mediated writing environments.

CONCLUSION

This study examined authorship consistency in EFL academic writing by conducting a forensic linguistic analysis of one questioned document against eight known samples from the same writer. The results showed that the participant's authentic texts maintained stable lexical, grammatical, syntactic, and stylistic habits, creating a clear linguistic profile. In contrast, the questioned document featured a much higher level of vocabulary sophistication, grammatical precision, sentence complexity, and formal academic tone, marking a sharp departure from the writer's baseline profile. Despite this linguistic divergence, analysis of the handwritten abstract reproduction task revealed strong conceptual continuity. The participant successfully reconstructed the principal components of the questioned document, including the research topic, objective, methodology, findings, and conclusion. These findings indicate substantial familiarity with the intellectual content represented in the questioned document and suggest that conceptual understanding may persist even when linguistic realization differs considerably.

The study's primary contribution is introducing the concept of conceptual-linguistic divergence. While traditional authorship studies look at the writer-text dynamic mostly through language similarities, this research proves that strong conceptual alignment and major linguistic differences can exist together in the same student authorship case. By establishing this framework, the study offers a new interpretative tool for investigating educational authorship within forensic linguistics. The research offers valuable practical insights for evaluating academic integrity and student authorship. In EFL contexts, a writer's linguistic output may fail to match their actual conceptual abilities. Because of this, looking strictly at language similarities risks oversimplifying the relationship between students and their texts, particularly in an educational landscape that is heavily shaped by AI and digital writing assistants.

Despite these contributions, the study is constrained by its focus on a single participant and one questioned document. To fully evaluate the scope and reliability of conceptual-linguistic divergence, future studies should utilize larger datasets, include multiple participants across various educational settings, and incorporate computational tools for authorship analysis.

REFERENCES

- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating? Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>
- Coulthard, M., & Johnson, A. (2010). *An Introduction to Forensic Linguistics: Language in Evidence* (2nd ed.). Routledge.
- Flowerdew, J. (2020). English for Research Publication Purposes and the Institutional Policies of Scholarly Journals. *Language Teaching*, 53(2), 180–197. <https://doi.org/10.1017/S0261444819000397>
- Fobbe, E. (2011). *Forensische Linguistik: Eine Einfuehrung*. Narr Francke Attempto Verlag.
- Hyland, K. (2003). *Second Language Writing*. Cambridge University Press.
- Hyland, K. (2019). *Second Language Writing* (2nd ed.). Cambridge University Press.
- Hyland, K., & Jiang, F. K. (2021). *Academic Discourse and Global Publishing: Disciplinary Persuasion in Changing Times*. Routledge.
- Juola, P. (2006). Authorship Attribution. *Foundations and Trends in Information Retrieval*, 1(3), 233–334. <https://doi.org/10.1561/15000000005>
- Kasneci, E., Sessler, K., Kuechemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Guennemann, S., Huellermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeiffer, F., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., & Kasneci, G. (2023). ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Lo, C. K. (2023). What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature. *Education Sciences*, 13(4), 410. <https://doi.org/10.3390/educsci13040410>
- McMenamin, G. R. (2002). *Forensic Linguistics: Advances in Forensic Stylistics*. CRC Press.
- Moorhouse, B. L., Yeo, M. A., & Wan, Y. S. (2023). Generative AI Tools and Educational Practice: Opportunities, Challenges, and Implications for Academic Writing. *Computers and Education: Artificial Intelligence*, 4, 100147.
- Nini, A. (2023). *A Theory of Linguistic Individuality for Authorship Analysis*. Cambridge University Press. <https://doi.org/10.1017/9781108974851>
- Perkins, M. (2023). Academic Integrity Considerations of AI Large Language Models in the Post-Pandemic Era. *Journal of University Teaching and Learning Practice*, 20(2), 1–12.
- Stamatatos, E. (2009). A Survey of Modern Authorship Attribution Methods. *Journal of the American Society for Information Science and Technology*, 60(3), 538–556. <https://doi.org/10.1002/asi.21001>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What If the Devil Is My Guardian Angel: ChatGPT as a Case Study of Using Chatbots in Education. *Smart Learning Environments*, 10(1), 15. <https://doi.org/10.1186/s40561-023-00237-x>