

Case Analyses on Implicit Academic Misconduct in Cross-Lingual Plagiarism

Cheng Zhang^{1,2, *}

¹Department of Basic Education, Jiangxi Police Institute, Nanchang, Jiangxi, China

²Faculty of Humanities and Social Sciences, City University of Macau, Macao SAR, China

*Corresponding Author: Cheng Zhang

ABSTRACT

Cross-lingual plagiarism-especially in the form of implicit academic misconduct-poses significant challenges to academic integrity, particularly in multilingual research contexts. This study explores the nuanced features of implicit academic misconduct within cross-lingual plagiarism, including syntactic interference, terminological inconsistency, inaccurate modifiers, and missing citations. Drawing on representative cases informed by real-world academic supervision, the paper illustrates how these subtle forms of misconduct often evade standard detection tools. The findings highlight the limitations of current plagiarism detection systems in capturing non-literal and multimodal content reuse. The study emphasizes the need for increased awareness, improved supervisory practices, and the development of more robust, cross-lingual-sensitive detection mechanisms to uphold academic integrity in increasingly globalized academic environments.

KEYWORDS

Cross-Lingual Plagiarism; Academic Misconduct; Plagiarism Detection.

1. INTRODUCTION

In recent years, the academic community has been increasingly confronted with cases of academic misconduct, including the falsification and plagiarism of dissertations. A particularly insidious form of this is the translation and publication of foreign academic papers, which is called cross-language plagiarism. According to statistics from CNKI, cross-language plagiarism has always been the hardest hit by plagiarism in China, but it is also one of the most difficult plagiarisms to detect[1]. Although numerous plagiarism detection tools are available, most rely on identifying text repetition and are not equipped to detect copied graphics, tables, formulas, citations, or multilingual content[2]. Consequently, academic misconduct that exhibits subtle characteristics-termed 'implicit academic misconduct'-often evades detection, complicating the assessment of academic integrity. This study gives the definition and features of implicit academic misconduct, and analyses in detail of a case study in order to improve the identification of implicit academic misconduct.

2. LITERATURE REVIEW

2.1. The Definition of Cross-Language Plagiarism

Cross-language plagiarism involves the republication of a translated version of a previously published academic paper or monograph. This practice is particularly prevalent in fields where there is a significant disparity between domestic and international research standards[3]. As a result,

researchers may be tempted to translate and publish foreign academic papers, which constitutes cross-language plagiarism. Studies on this topic conducted abroad have typically focused on plagiarism detection between languages within the same family, which share common word formations, grammatical structures, and cultural backgrounds. In contrast, domestic research on plagiarism has predominantly addressed plagiarism within Chinese-language articles, with limited attention to cross-language issues[4]. The few studies that exist are mainly within the field of linguistics and aim to facilitate linguistic comparisons rather than detect plagiarism.

2.2. The Features in Cross-Lingual Plagiarism

Cross-language plagiarism can be identified by several distinct features:

- 1) The text may lack fluency, indicating that it has been translated rather than originally written in the target language.
- 2) Expressions in the text may not conform to the norms and conventions of the target language.
- 3) The content may require careful interpretation to be understood, reflecting the challenges of translating complex academic ideas.
- 4) Sentence structures may exhibit noticeable influences from the source language.
- 5) Modifiers such as adjectives and adverbs may be used inaccurately or imprecisely.
- 6) Terminology may be inconsistent, with different terms referring to the same concept, or professional terms may be non-standardized.
- 7) References may be predominantly in the source language, with few or none in the target language.

2.3. Types of Cross-language Plagiarism

Cross-language plagiarism takes the following forms:

- (1)Direct Copying. This involves copying content directly from the source document without modification. This form of plagiarism is straightforward, with little to no changes made to the copied content. The copied sections may be almost identical to the source, with only minor alterations to words or punctuation.
- (2)Sentence synonym conversion. This involves making modifications to the text by replacing words with synonyms, changing sentence structures, or rearranging the order of sentences. This type of plagiarism can include abbreviating, expanding, or splitting sentences to disguise the copied content.
- (3)Language Transformation. This involves translating content from the source document into another language to create a plagiarized work. For example, an English document may be translated into Chinese, or vice versa, to obscure the original source.

3. A CASE STUDY

To further illustrate the characteristics of implicit academic misconduct in cross-lingual plagiarism, this section presents three representative cases. Each case reflects typical features outlined in Section 2.2, such as syntactic interference, terminological inconsistency, inaccurate use of modifiers, and lack of proper citation. These scenarios are grounded in real-world academic supervision and provide practical relevance.

Case 1: Terminological Inconsistency and Structural Interference

An undergraduate student submitted a final thesis exploring the Theme-Rheme structure in Chinese media discourse, which focused on the application of Systemic Functional Linguistics (SFL).

Although the initial plagiarism detection score was only 13.7%, the writing style raised significant concerns. The use of specialized terminology was highly inconsistent—for instance, the term “Theme-Rheme” was rendered alternately as “主题-述题”, “主位-述位”, and “主题结构” throughout the text. This variation suggested a lack of terminological understanding and a probable reliance on multiple machine translation outputs.

Moreover, the syntactic structure of many sentences appeared to be influenced by English grammar patterns. One notable example read: “话语单位的组织是由主题的突出所决定的, 这种主题的突出是信息上被标记的.” This phrasing, while technically grammatical, was highly unnatural and echoed the structure of an English sentence found in Halliday’s original 1994 work.

Further manual comparison revealed that over 50% of the content was translated from a peer-reviewed English article published in 1998. The student had failed to cite the source or include any English-language reference in the bibliography. The case was identified as implicit cross-lingual plagiarism.

Case 2: Paragraph-Level Translation and Expression Inaccuracy

Another example involved a student writing on Second Language Acquisition (SLA), specifically on Krashen’s input hypothesis. The thesis was well-organized and featured smooth paragraph transitions. However, the supervisor noted that certain segments seemed unusually fluent and thematically cohesive compared to the student’s prior coursework.

Upon closer analysis, it became apparent that two paragraphs had been closely translated from an English-language research paper. The translated text retained the structure, argument flow, and even examples used in the original English article. Below is a comparison between a section of the original article and the student’s Chinese text:

Table 1. Source Text and Translation Comparison

Original English Text	Translated Chinese text in Student Thesis
"Language input enhances the learner’s ability to construct output hypotheses through comprehension mechanism."	“语言输入通过理解机制增强了学习者构建输出假设的能力。”
“When learners are exposed to rich input, their capacity for producing target language output increases.”	“当学习者接触到丰富的输入时, 他们产生目标语输出的能力会增强。”

Although the translation was relatively fluent, it still contained expressions that were atypical in academic Chinese writing. For instance, “输出假设” and “理解机制” were not properly contextualized within the Chinese text, and their meanings remained vague. Furthermore, the reference list contained no mention of the original article or any related English publications.

The plagiarism report shows a low similarity score (12.6%) and failed to detect the issue due to the cross-lingual nature of the misconduct. Nevertheless, the case clearly aligned with the characteristics of implicit academic plagiarism.

Case 3: Reuse of Figures and Equations Without Attribution

A postgraduate student submitted a thesis on translation alignment algorithms used in computer-assisted terminology extraction. The thesis included several sophisticated figures, flowcharts, and mathematical expressions. While the body text passed plagiarism checks with a score of 10.2%, the visual materials raised suspicion due to their professional layout and technical accuracy, which seemed inconsistent with the rest of the writing.

Upon investigation, it was found that the figures had been taken from an English-language journal article published in a computational linguistics journal. The student had translated the figure captions into Chinese but had not changed the structure or content of the graphics. Furthermore, these figures were not cited or credited anywhere in the text or bibliography.

This case exemplifies non-textual cross-lingual plagiarism, where the misconduct lies not in linguistic reuse but in the unauthorized appropriation of visual or formulaic academic content. This aligns with the challenges discussed in [5], who argue that current detection software cannot effectively capture multimodal plagiarism, including graphics, tables, and equations. The case highlights the importance of removing improperly attributed visual content and ensuring originality in both textual and non-textual academic materials.

These three cases demonstrate the complexity of identifying and addressing cross-lingual plagiarism, especially when it takes subtle or non-literal forms. Whether through inconsistent terminology, structure-preserving translation [6], or the uncredited use of non-textual materials, implicit academic misconduct undermines the integrity of scholarly communication. This highlights the need for improved supervisory practices and a broader understanding of plagiarism that goes beyond textual repetition alone.

4. CONCLUSION

In recent years, with the rapid development of network technology, it has become more and more convenient for people to obtain and share information. Digital files from the network can be easily copied, and anyone can easily obtain information from other people's files, thus constituting plagiarism. The newly issued Law of Academic Degrees on April 26 in 2024 has strengthened the regulation of academic misconduct, establishing legal foundations for improving education quality and fostering innovation. This underscores the necessity of research into anti-plagiarism measures to protect intellectual property and maintain a harmonious academic environment. Anti-plagiarism efforts are not only essential for protecting intellectual property rights but also for fostering an environment of academic integrity and innovation.

REFERENCES

- [1] Alzahrani, S. M., Salim, N., & Abraham, A. (2012). Understanding plagiarism linguistic patterns, textual features, and detection methods. *IEEE Transactions on Systems, Man and Cybernetics. Part C, Applications and Reviews*, 42(2), 133–149. <https://doi.org/10.1109/tsmcc.2011.2134847>
- [2] Alzahrani, S., & Aljuaid, H. (2022). Identifying cross-lingual plagiarism using rich semantic features and deep neural networks: A study on Arabic-English plagiarism cases. *Journal of King Saud University. Computer and Information Sciences/Mağalaġ Ğam'aġ Al-malik Saud : Ūlm Al-ħasib Wa Al-ma'lumat*, 34(4), 1110–1123.
- [3] Asghari, H., Fatemi, O., Mohtaj, S., Faili, H., & Rosso, P. (2019). On the use of word embedding for cross language plagiarism detection. *Intelligent Data Analysis*, 23(3), 661–680. <https://doi.org/10.3233/ida-183985>
- [4] Haneef, I., Nawab, R. M. A., Munir, E. U., & Bajwa, I. S. (2019). Design and development of a large Cross-Lingual Plagiarism Corpus for Urdu-English Language pair. *Scientific Programming*, 2019, 1–11. <https://doi.org/10.1155/2019/2962040>
- [5] Hourrane, O., & Benlahmar, E. H. (2022). Graph transformer for cross-lingual plagiarism detection. *IAES International Journal of Artificial Intelligence*, 11(3), 905. <https://doi.org/10.11591/ijai.v11.i3.pp905-915>
- [6] Lijuan Qi & Jingyi Ji.(2021) .Identification and prevention of inapparent academic misconducts: Cross-language plagiarism and ghostwriting. *Research on Chinese science and technology journals*, 32(11).