Tracing Plagiarism: Linguistic Indicators and Gender-Based Patterns in Forensic Linguistics

Sri. Praveenkumar
Research Scholar, Department of Linguistics
Bharathiar University, Coimbatore

Dr. N. Vijayan
Assistant Professor, Department of Linguistics
Bharathiar University, Coimbatore

Abstract

"Plagiarism represents a spectrum of behaviors from copying text verbatim to closely paraphrasing without credit, challenging academic integrity. "Neumann (2002)This can take the form of copying text, paraphrasing without proper citation, or using someone else's images, videos, or other forms of media without proper attribution. Plagiarism is considered unethical and can have serious consequences in educational, professional, and legal contexts. Forensic Linguistic analysis of plagiarism involves identifying and comparing the writing style, language, and other linguistic features of a suspected plagiarized text to the original source material. This study investigates the linguistic evidences to identify the usage of language among male and femalein plagiarised.

Introduction

"Plagiarism is the reproduction or appropriation of someone else's work without proper acknowledgment, resulting in a failure to attribute intellectual property to its rightful owner."Howard, R. M. (1999). There are several types of plagiarism Copying verbatim text without citation, paraphrasing someone else's work without proper citation, submitting someone else's work as your own, Self-plagiarism, which is reusing your own work without proper citation, Patchwork plagiarism, which is combining various sources without proper citation. All of these types of plagiarism are considered unethical and can have serious consequences.

When it comes to point of detection of plagiarism there aremany tools available that use algorithms to compare texts and identify similarities. These tools can be useful in detecting plagiarism, but they are not foolproof and can sometimes produce false positives or miss cases of plagiarism.Linguistic analysis can be used to detect plagiarism by comparing the writing style, vocabulary, and sentence structure of the suspected plagiarized text with the original text.

Gender and language are two important areas of research in linguistics that have been studied extensively over the past few decades. While gender refers to the social and cultural roles and expectations associated with being male or female, language is the medium through which we communicate and express our ideas, thoughts, and emotions. The study of gender and language seeks to understand the ways in which language use reflects and reinforces gender roles and stereotypes, as well as how gender affects the ways in which individuals communicate and are perceived by others. This study investigates about the language use of male and female in plagiarized text.

Aim

To investigate the linguistic markers that facilitate plagiarism and to explore gender-based differences in the patterns and strategies of plagiarism through a forensic linguistic lens

Objectives
- To identify the linguistic characteristics of plagiarism
- To identify the Gender differences in plagiarized text

Data Source

A small corpus of empirical data was used, it consist of article about Mahatma Gandhi ji had an average of 300 words was circulated among the students of second year linguistics postgraduate students. They are instructed by the researcher to rephrase the article in their own words without any limitations. The research collected the data and used convenient sampling method for analysis.

Methodology Used

The researcher used Discourse analysis methods for analysis. The researcher adopted qualitative method for identifying plagiarism and quantitative method for analysing the differences between the male and female in plagiarized text.

Analysis

Syntactic Rearrangement

Syntactic rearrangement in plagiarism refers to altering the structure of sentences in an attempt to disguise copied content. This tactic is often used by plagiarists to make the borrowed text appear original while retaining the original meaning and ideas. It is a significant linguistic marker in forensic linguistics for identifying potential cases of plagiarism.

Source Text

1. Mahatma Gandhi, born on October 2, 1869, is celebrated as a key figure in India's struggle for independence. Known as the "Father of the Nation," Gandhi championed the principles of non-violence and truth.
2. His approach, called Satyagraha, involved peaceful protests and civil disobedience to challenge British rule.

Plagiarised Text

Born on October 2, 1869, Mahatma Gandhi is revered as a pivotal player in India's fight for freedom. Gandhi, who was referred to as the "Father of the Nation," promoted the values of truth and non-violence.

His strategy, known as satyagraha, challenged British rule through nonviolent demonstrations and civil disobedience.

In the about mentioned extract the plagiarist changed the syntactic order of the paragraph. In the original, "Mahatma Gandhi, born on October 2, 1869," is the opening phrase. In the plagiarized

text, the order is reversed to "Born on October 2, 1869, Mahatma Gandhi," keeping the same information but changing the sentence structure.

The original sentence, "His approach, called Satyagraha, involved peaceful protests and civil disobedience to challenge British rule," is restructured to "His strategy, known as Satyagraha, challenged British rule through nonviolent demonstrations and civil disobedience."The order of phrases has been changed ("to challenge British rule" becomes "challenged British rule through"). The changes retain the original meaning but alter the wording and structure to make the text appear less identical while still conveying the same idea, which is a common strategy used in plagiarism.

Deleting Specifications

Deleting specifications in plagiarism refers to the act of omitting or removing specific details, data, or references from the original source to make the copied material appear more generalized, often to obscure its origin. While the core ideas or concepts may be retained, the absence of specific details can mislead the reader into thinking that the work is original.

Source Text

One of his most notable campaigns was the Salt March of 1930, where he walked over 240 miles to protest the British monopoly on salt. This act of peaceful resistance mobilized millions and drew international attention to India's quest for freedom.

Mahatma Gandhi's efforts against 'untouchability' while Gandhi was in Yerwada Jail Pune, India on September 16, 1932. The hunger strike was a protest against the British government's proposal to separate India's electorate by caste.

Plagiarized Text

"His notable campaigns was a march where he walked to protest the monopoly on salt. This act of peaceful resistance mobilized many and drew attention to quest for freedom."

Mahatma Gandhi's efforts against 'untouchability' while Gandhi was in Yerwada Jail, where he went on a hunger strike.

In addition to generalizing the instances, the plagiarists removed the date, time, and location details from the above. The length of the march and the date of the salt sathygraha were concealed by the copied information. In example 4, the information about when Gandhji initiated the hunger strike and its cause was concealed.

Adding Specifications

Adding specifications in plagiarism refers to the act of introducing additional, often irrelevant or exaggerated, details to the original content in an attempt to mask the plagiarism or make the text appear more original. While this strategy may alter the surface structure of the copied material, the core idea or argument remains plagiarized.

Source Text

Gandhi also worked to improve social conditions, fighting against untouchability and promoting unity among different communities. His commitment to simplicity, truth, and non-violence made a significant impact, both in India and globally

Plagiarised Text

"Gandhi also worked on improving societal conditions, challenging social inequalities and advocating for harmony among diverse groups. His dedication to basic values like honesty,

simplicity, and peaceful resistance had a profound effect, not only in India but also on a global scale."

Compared to the original, the plagiarized material is less specific and broader due to the addition of certain extra specifications. More generic phrases like "social inequalities" and "harmony among diverse groups" have taken the place of the original particular like "untouchability" and "unity among different communities.""Commitment to simplicity, truth, and non-violence" =>"Dedication to basic values like honesty, simplicity, and peaceful resistance.""Honesty" and "peaceful resistance" take the place of the first ideals of "truth" and "non-violence." This gradually changes the focus, making it more general and less specific. Instead of using the exact words used in the original, the phrase "basic values like honesty" is somewhat broadened to generalize Gandhi's values.

Verbatim

Verbatim plagiarism refers to the act of copying someone else's work word-for-word without proper citation or attribution. It is one of the most direct and obvious forms of plagiarism, where the plagiarist uses the exact language, sentence structure, or phrasing of the original source, often with minimal or no modification.

Source Text

His idea of 'Ahimsa,' which means not hurting anyone, was widely admired and followed by many influential people worldwide.

Plagiarized Text

He popularized the concept of 'Ahimsa,' which means not harming anyone, and many powerful individuals around the world adopted it.

Usage of Pronoun

The use of personal pronouns in plagiarism can be a key factor in distinguishing between original work and plagiarized content, particularly when it comes to the ownership of ideas or the structure of the text. Personal pronouns (e.g., I, we, he, she, they, and you) are often used to attribute the perspective or voice of the writer, and their usage—or lack thereof—can offer insight into how an individual approaches their writing or copying.

Source Text

Mahatma Gandhi, born on October 2, 1869, is celebrated as a key figure in India's struggle for independence. Known as the "Father of the Nation," Mahatma Gandhi championed the principles of non-violence and truth. Mahatma Gandhi approach, called Satyagraha, involved peaceful protests and civil disobedience to challenge British rule. Mahatma Gandhi made significant contributions to various social issues. Mahatma Gandhi's efforts against 'untouchability' while Gandhi was in Yerwada Jail, where he went on a hunger strike.

Plagiarized Text

Mahatma Gandhi, born on October 2, 1869, is celebrated as a key figure in India's struggle for independence. Known as the "Father of the Nation," hehampioned the principles of non-violence and truth. His approach, called Satyagraha, involved peaceful protests and civil disobedience to challenge British rule. Hemade significant contributions to various social issues. His efforts against 'untouchability' while he was in Yerwada Jail, where he went on a hunger strike.

Usage of Synonyms

The usage of synonyms in plagiarism refers to the strategy where a plagiarist replaces words or phrases in the original text with their synonyms to make the content appear more original. This is a common form of paraphrasing, though it can still be considered plagiarism if the overall structure, ideas, and meaning remain intact without proper attribution to the original source.

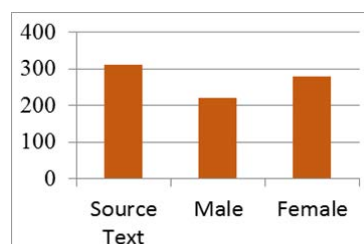| S.No	Actual words in source text	Change of Synonyms in Plagiarized text
1	Social conditions	Societal conditions
2	Fighting against untouchability	Challenging social inequalities
3	Promoting unity among different communities	Advocating for harmony among diverse groups
4	Pivotal player	key figure
5	Fight for freedom	struggle for independence
6	Promoted	championed
7	Approach	strategy
8	Involved	challenged
9	Peaceful protests	nonviolent demonstrations

Substitution of synonyms is major linguistic character in plagiarized text her the plagiarist employed substitution throughout the passages to give the exact meaning of the essay whithout any change.

Gender Differences in Plagiarized Text

The researcher conducted quantitative analysis on plagiarized text to identify gender differences in the context of authorship identification involves examining various linguistic features that may differ between male and female authors. While authorship identification can be quite complex, focusing on gender differences within plagiarized texts involves analysing linguistic markers that might be indicative of the author's genderJuola, Patrick. (2008): 233-334.. Frequency analysis is employed by the researcher to identify the gender differences in plagiarism.Harmon, R. (2004)The researcher identified differences in word count, paragraph, sentences count, intentional spelling change, numerals etc. here going to see on detail,

Word Count

Word count plays an important role in authorship identification by providing quantitative data that can help differentiate between various writing styles.Different authors may have distinctive writing habits that affect their overall word count in a given text. Some authors may tend to write longer sentences or use more descriptive language, resulting in higher word counts, while others may write concisely, resulting in shorter texts. By analysing the average word count across various texts written by the same author, researchers can identify patterns that can help distinguish authorship.
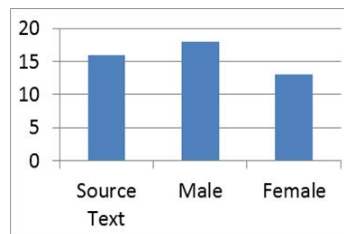
The above clearly explains that men use fewer words when they plagiarizing the article, women uses more emphasis, references to prove their argument which results in usage of more words when compared to male.

Sentence Count

Sentence count plays a crucial role in authorship identification, particularly in distinguishing between male and female writing styles. It involves analysing the number of sentences within a given text and can provide insight into the structure and organization of writing, which may vary between male and female authors.

Results of the chart declares that Men use more sentences the researchsuggest that female writers often produce shorter sentences compared to their male counterparts. This is due to a preference for concise and clear communication, which is common in many genres of writing by women. Female writers may also use simpler, more direct sentence structures.
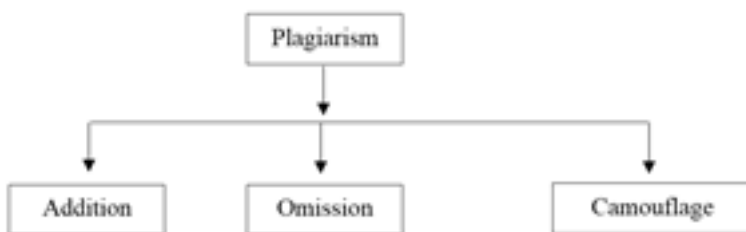
Major Differences Found in Comparison of Male and Female

S.No	Characters	Male	Female
1	Change in spelling	High	Low
2	Adding information	Low	High
3	Omitting information	High	Low
4	Using synonyms	Low	High
5	Using hyponyms	Low	High
6	Using honorary titles	Low	High
7	Concealment of information	High	Low
8	Using adversative conjunction	Low	High
9	Usage of punctuations	low	high

Findings

The research had two basic objectives which is what the Linguistic markers which involve in Plagiarism? And what are the differences between the male and female in plagiarizing the text, with the help of discourse analysis the researcher identified that there only three basic deceptive characters in the plagiarism,

Addition includes the addition of Honorifics, adjective, articles and references to prove that their writing is original, Omission includes the deletion of honorifics, titles, information about events etc., camouflaging in plagiarism have a deep insight where the plagiarist uses synonyms of the same semantic quality where employed, the plagiarist uses pronouns as camouflaging technique to prove their writing as the real one. The information's were camouflaged in the plagiarized text.

The Forensic Authorship Identification method is used to identify the differences between male and female plagiarists. It gave deep insight that men uses less word count when compared to female since they are using more emphasisers to prove their argument. Women uses more honorifics,

usage of adversative conjunction and also the usage of punctuations. Thus this article provides the linguistic characteristics of plagiarism and also furnishes the differences between the characteristics of male and female in plagiarized text.

References

1. Juola, Patrick. "Authorship attribution." Foundations and Trends® in Information Retrieval 1.3 (2008): 233-334.
2. Harmon, R. (2004). Frequency Analysis in Linguistics and Forensic Text Examination. Journal of Computational Linguistics, 16(3), 234-245.
3. Howard, Rebecca Moore. "Sexuality, textuality: The cultural work of plagiarism." College English 62.4 (2000): 473-491.
4. Scanlon, Patrick M., and David R. Neumann. "Internet plagiarism among college students." Journal of College Student Development 43.3 (2002): 374-385.
5. Sousa-Silva, Rui. "Investigating academic plagiarism: A forensic linguistics approach to plagiarism detection." International Journal for Educational Integrity 10.1 (2014).
6. Sousa Silva, Rui, Tim Grant, and Belinda Maia. ""I didn't mean to steal someone else's words!: a forensic linguistic approach to detecting intentional plagiarism." 4th International Plagiarism Conference. 2010.