

PHILOLOGICAL SCIENCES

UDC: 81'367'42:004.8

INSIGHTS INTO HUMAN VS AI DISCOURSE PRODUCTION

Fedoriv Yaroslava Romanivna,

Candidate of Philology, Associate Professor

Pirozhenko Iryna Dmytrivna,

Shuhai Alla Yuriivna,

Senior Lecturers

The National University of “Kyiv-Mohyla Academy”

Kyiv, Ukraine

Summary. Distinguishing between non-native human-written English text and AI-generated text presents a notable challenge due to their shared characteristics that serve as means of communication. This paper presents several indicators that can assist in forming a judicious assessment, such as grammar, syntax, unnatural language use, cohesion, context and coherence, consistency and response patterns, background knowledge, cultural awareness, and credibility.

Introduction Differentiating between human-created and AI-generated texts has become increasingly significant due to advancements in text-processing technologies and the integration of artificial intelligence in various communication domains such as education, journalism, and content development. Both forms share similarities in conveying information and adhering to grammatical rules to communicate meaningful messages, emotions, stories, and diverse responses.

This research **aims** to explore linguistic tools and identify features distinguishing human and AI discourse in order to preserve human creativity in communication, ensure fair text attribution, and gain insights into the implications for EFL proficiency assessment.

I. Linguistic Characteristics in Human-Created Discourse

Human-created discourse relies on fundamental **grammatical** and **syntactical**

aspects for effective communication in written and spoken language. Human-produced texts demonstrate understanding of the aspectual properties of verbs and functional organisation of clauses in linguistic expressions [5, p. 57–68]. However, non-native human-written texts may exhibit grammatical errors, unconventional word order, or inconsistent verb forms.

Another aspect to consider is **unnatural language use**, which arises when non-native human-written texts display linguistic behaviours or peculiar ways of expression. Vocabulary usage can vary in non-native human-written texts due to language background, resulting in language errors specific to the writer's native tongue [7; 12, p. 215-222; 13]. A divergence in linguistic choices can lead to misunderstandings and difficulties in conveying intended meanings, affecting the clarity and effectiveness of communication.

Additionally, non-native human-written texts may exhibit inconsistencies in **style**, with the use of figurative language as creativity markers [9, p. 253-256].

Response patterns. Consistency is defined as “the orderly treatment of a set of linked elements” [4]. In non-native human-written texts response patterns may vary in complexity and clarity, impacting how information is presented and perceived. Understanding and analysing these response patterns are essential, especially in the area of English as a foreign language.

Another crucial aspect of meaningful communication is **cohesion** which ensures a sense of unity and **coherence** in written and spoken content. Human cohesion results from a complex cognitive process influenced by linguistic competence and personal experiences. Humans use linguistic devices such as pronouns, connectives, conjunctions, and lexical repetition to establish logical connections between ideas and ensure smooth transitions [3; 10].

Text cohesion includes not only cohesive devices within the text, but also those that relate text with the situation [2].

Proficient language users possess a deeper understanding that goes beyond language skills, while non-native English users may encounter coherence issues due to cultural or knowledge gaps [8, p. 200-220]. In some instances, non-native speakers

may endure miscommunication even with a strong command of the target language, which is known as “pragmatic failure” [1].

Knowledge and cultural awareness play a significant role in determining the types of content. This implies the cultural specificity of human language, conducted mostly in modern, multiethnic, industrialised societies [11, p. 43-59]. Non-native human-written texts may show gaps in cultural knowledge but provide insights specific to the writer's culture or background.

Proper referencing in academic papers is crucial to ensure the **credibility and integrity** of research, giving credit to original authors and allowing readers to access and explore cited sources for transparency and further research in the corresponding field.

II. AI-Generated Discourse Characteristics

The analysis of AI-generated texts reveals several distinctive features.

AI adheres to **grammatical rules**, maintaining consistent **syntax** and parallel structures by using statistical patterns learned from extensive training data, resulting in precision and coherence [5]. Notably, AI employs contracted verb forms, contributing to its strikingly rhythmic and monotonous tone, which can be a potential indicator of its artificial origin. In terms of language usage, AI-generated texts often exhibit an excessively formal or rigid style, lacking colloquial expressions and natural linguistic patterns found in human-created content, further setting them apart [7; 12].

Unlike human-created texts, AI **cohesion** relies on statistical patterns derived from native English speakers' discourse, enabling it to produce texts with some level of coherence. However, AI cohesion has limitations, as it may fail to maintain consistent focus in long-term narratives or display deep contextual understanding [3; 10].

In distinguishing between human-written and AI-generated content, factors like **coherence, consistency in response patterns**, and background knowledge become essential. The presence of consistent patterns in style, vocabulary, or response structure may indicate AI involvement.

AI's lack of **contextual knowledge**, personal experiences, and cultural understanding can lead to occasional errors or contextually inappropriate responses, underscoring the role of background knowledge and cultural awareness in differentiation [8]. For instance, AI-generated texts may display a lack of updated information or knowledge of current events [13]. Additionally, AI may offer "miscitations," providing inaccurate or erroneous references, which compromises the integrity and trustworthiness of the content [6].

In light of these observations, distinguishing between human and AI discourse is crucial for fostering responsible and reliable communication. Identifying patterns in style, coherence, and contextual appropriateness can aid in determining the text's origin. As AI continues to evolve, understanding these features will play a pivotal role in maintaining transparency and academic integrity in the digital era.

Conclusion The analysis of human-created and AI-generated texts highlights certain features differentiating them. Key aspects include grammar, syntax, unnatural language use, cohesion, context, coherence, consistency, response patterns, background knowledge, cultural awareness, and credibility. AI text is characterised by a rhythmic, monotonous tone, contracted verb forms, and limited cohesion. Human texts display a unique voice, creative patterns, and colloquial expressions. AI-generated content lacks contextual awareness and accurate responses, often using fictional references. Recognising these issues is essential for maintaining transparency and academic integrity when encountering AI-generated texts, ensuring responsible communication in the digital era and upholding ethical use of language technologies.

REFERENCES

1. BAO, X. (2020). Pragmatic Failure in Cross-Cultural Communication. *Cross-Cultural Communication*. <http://cscanada.net/index.php/ccc/article/view/11866/11796>
2. Delu, Z. (2005). Form and Meaning in Text Cohesion. *Journal of foreign languages*. <https://www.semanticscholar.org/paper/Form-and-Meaning-in-Text->

3. Er, Z. (2004). Cohesion by Means of Words in an Article. <https://www.semanticscholar.org/paper/Cohesion-by-Means-of-Words-in-an-Article-Er/db8caaf14d31e5be0a6f402f9a55f1de689d7d9f>
4. Farkas, D. K. (1985). The Concept of Consistency in Writing and Editing. *Journal of Technical Writing and Communication*. <https://doi.org/10.2190/T6EM-UTT0-EL6J-59N9>
5. Fries, C. C. (1954). Meaning and Linguistic Analysis. *Language*, 30(1), 57–68. <https://doi.org/10.2307/410220>
6. Harzing, W. (2002). Are our referencing errors undermining our scholarship and credibility? The case of expatriate failure rates. *Journal of Organizational Behavior*, 23(1), 127-148. Copyright © 2002 John Wiley & Sons, Ltd. <https://doi.org/10.1002/job.125>
7. Lewis, K. (2020). False Friends in Terminology. *Semantic Scholar*. https://pdfs.semanticscholar.org/5858/42caffb6c73e6ec569311b402f50d3d5ce39.pdf?_gl=1*iiuj0*_ga*MTQ4Nzc2NjU3MS4xNjg5NjA3NDAx*_ga_H7P4ZT52H5*MTY4OTYwNzQwMC4xLjEuMTY4OTYwNzcyOC4zNy4wLjA.
8. Matthews, R. J. (2006). Knowledge of Language and Linguistic Competence. *Philosophical Issues*, 16, 200–220. <http://www.jstor.org/stable/27749865>
9. Price, J. T. (1974). Linguistic Competence and Metaphorical Use. *Foundations of Language*, 11(2), 253–256. <http://www.jstor.org/stable/25000771>
10. Qing-yuan, L. (2000). On the Function of Cohesive Devices in a Text A Model of Text Analysis. *Journal of Chenzhou Teachers College*. <https://www.semanticscholar.org/paper/On-the-Function-of-Cohesive-Devices-in-a-Text-%E2%80%94Qing-yuan/d8a2c96a310de5891fcb6f31b2da0c2e68554c8b>
11. Singh, R., Lele, J., & Martohardjono, G. (1988). Communication in a multilingual society: Some missed opportunities. *Language in Society*, 17(1), 43-59. doi:10.1017/S0047404500012586

12. Topalova, A. (2010). 'False friends' in translation work: An empirical study. *Perspectives*, 4:2, 215-222, DOI: 10.1080/0907676X.1996.9961288

13. Veisbergs, A. (1996). False Friends Dictionaries: A Tool for Translators or Learners or Both. <https://semanticscholar.org/paper/f531d3b9bb81a6f90acc45cf1065368ec4a273ee>