

# A Comparative Study on the Meta-functions of Artificial Intelligence and Human Discourse Markers

SHI Ren-wen

University of Shanghai for Science and Technology, Shanghai, China

This study explores the use of discourse markers in AI-generated and human-generated texts, based on Halliday's metafunctional theory. By analyzing 20 Shanghai college entrance exam essays generated by Baidu's AI model Wenxinyiyan and comparing them with 20 human-written essays, it is found that both share similar use of discourse markers. However, AI tends to use transitional and topic-follow-up markers more frequently, while human texts show a higher usage of conditional and causal markers. These findings suggest AI's capacity for maintaining discourse coherence but highlight its limitations in more complex reasoning.

*Keywords:* discourse, metafunctions, discourse markers, AI

## Introduction

Artificial Intelligence (AI) has made significant progress in natural language processing (NLP), particularly in generating coherent texts that mimic human language. Discourse markers, crucial for logical connections and coherence, have become an important focus in understanding AI's capabilities compared to human writing. This paper examines the differences in discourse marker usage between AI and humans, providing insights into areas where AI-generated texts still fall short of human fluency and emotional depth.

## Literature Review

Numerous studies have explored AI-generated text, particularly focusing on how AI compares to human creativity. Clerwall (2014) pioneered the comparison of machine-generated and human-generated texts. Later research, such as Gunser (2021) and Kobis & Mossink (2021), found that AI can produce highly similar texts to human writing but tends to rely on repetitive patterns. In contrast, human texts demonstrate flexibility in word choice and structure. Domestic scholars like Zhu Junhui et al. (2024) revealed that AI-generated language lacks the richness and diversity found in human discourse. This research builds on these studies by focusing on a specific linguistic phenomenon: discourse markers.

## Research Methods

The study uses a corpus consisting of 20 AI-generated texts from Wenxinyiyan and 20 human-written texts from the same essay prompts. Discourse markers were manually annotated and classified into nine

categories, including transitional, conditional, and causal markers. Quantitative analysis was conducted using AntConc, while a t-test was applied to assess statistical differences in marker frequency.

### Discourse Markers in AI-generated Texts

In this study, the discourse markers in the 20 college entrance examination essays generated by Wenxinyiyan were carefully classified and their usage frequencies were counted. The specific results are shown in Table 1.

Table 1  
*Frequency Distribution Table of Discourse Markers in AI-generated Texts*

| Types            | Frequency |
|------------------|-----------|
| Transition       | 116       |
| Topic follow-up  | 52        |
| uncertainty      | 30        |
| Example          | 20        |
| Conditional      | 17        |
| Summary          | 17        |
| Ordinal          | 12        |
| Cause and Effect | 11        |
| Clarification    | 5         |

The data from Table 1 shows that Wenxinyiyan frequently uses transitional discourse markers, such as “but” and “however”, to ensure coherence and logical flow. Topic-follow-up markers, like “in addition” and “next”, help AI expand topics, while uncertainty markers, such as “maybe” and “probably”, reflect caution in handling speculative information.

Other markers, like “for example”, provide examples, and conditional markers like “if” introduce hypothetical scenarios. Summary markers and ordinal markers help structure the text, though AI’s reliance on a fixed framework can sometimes result in rigid and less flexible writing compared to humans, who tend to use more varied and personalized expressions.

Overall, Wenxinyiyan demonstrates clear logic and organization through discourse markers, but AI-generated texts can be stereotyped and lack emotional depth, signaling an area for improvement in AI-generated language.

### Discourse Markers in Human-generated Text

The frequency of use of discourse markers in the 20 college entrance examination essays written by humans is shown in Table 2.

Table 2  
*Frequency Distribution Table of Discourse Markers in Human-generated Texts*

| Type             | Frequency |
|------------------|-----------|
| Transition       | 140       |
| Conditional      | 49        |
| Topic follow-up  | 28        |
| Cause and effect | 21        |
| Uncertainty      | 17        |

|               |    |
|---------------|----|
| Clarification | 12 |
| Example       | 4  |

Table 2 shows that transitional discourse markers also play a key role in human writing, helping maintain logical flow and structure. Markers like “but” and “however” introduce contrasting ideas, enhancing argumentation and reflecting the complexity of human thinking.

Humans use conditional markers, such as “if”, more frequently, allowing for hypothetical reasoning, while causal markers like “because” show clear logical connections. Although topic-follow-up and uncertainty markers are used less often, they still help maintain smooth logic and objectivity in uncertain statements.

Fact-clarifying markers like “in other words” are less common but important for clarifying complex ideas, guiding readers to deeper understanding. Humans also tend to use detailed descriptions instead of simple example markers like “for example” making their writing more vivid and persuasive.

Overall, human authors use discourse markers flexibly, especially for logical reasoning, condition setting, and clarifying information, which adds depth and richness to their writing.

### Meta-function Comparative Analysis

When comparing AI and humans in discourse marker use, based on Halliday’s meta-functional theory, AI tends to use transitional markers less than humans, showing limitations in complex logical transitions. However, AI frequently uses example markers to enhance clarity, though these are often repetitive and lack flexibility or emotional depth, making AI-generated text feel mechanical. Humans, on the other hand, adapt language based on context and emotional needs, making their writing more personalized and diverse.

AI’s use of uncertainty markers, like “maybe”, reflects caution and objectivity, but lacks emotional resonance. In contrast, human authors use fact-clarifying markers, like “actually”, to build trust and clarify ideas, fostering deeper reader engagement.

Overall, AI tends to remain neutral and avoid errors through uncertainty, while humans clarify facts to build consensus. AI maintains coherence through topic follow-up markers, but its limited use of conditional markers affects logical rigor. Despite similarities in overall marker use, AI still lags behind humans in handling more nuanced and flexible communication. Statistical tests show no significant overall difference in marker frequency, but AI’s limitations in specific categories warrant further study.

### Results

This study compares the types and frequencies of discourse markers in AI and human-generated texts, addressing three main questions:

(1). While there are differences in specific marker categories between AI (Wenxinyiyan) and humans, overall, the difference in marker usage is not significant ( $t=-0.125$ ,  $p=0.903$ ).

(2). Discourse markers serve key meta-functions in both AI and human discourse. Transitional markers are the most common, supporting coherence, while topic follow-up and conditional markers show AI’s ability to advance topics and humans’ strength in logical reasoning. Causal, uncertainty, and clarification markers also play important roles in both.

(3). Wenxinyiyan frequently uses transition and topic follow-up markers but shows lower use of conditional and causal markers, indicating a need to improve its logical reasoning capabilities. Enhancing the AI's algorithm and training data diversity could address these gaps.

### Conclusion

This research shows that AI (Wenxinyiyan) closely resembles human use of discourse markers, particularly in transition and topic follow-up markers. While there are some category differences, overall, the frequency of marker usage between AI and humans is not significantly different. These findings offer insights for improving the naturalness and logic of AI-generated text. Future research could expand the sample size, control more variables, and explore additional generative tasks to better understand the similarities and differences between AI and human discourse.

### References

- Clerwall, C. (2014). Enter the robot journalist. *Journalism Practice*, 8(5), 519-531.
- Gunser, V. E., Gottschling, S., Brucker, B., et al. (2021). Can users distinguish narrative texts written by an artificial intelligence writing tool from purely human text? *HCI International 2021 Posters*, 520-527.
- Halliday, M. A. K. (2004). *An introduction to functional grammar*. London: Edward Arnold.
- Köbis, N., & Mossink, L. D. (2021). Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry. *Computers in Human Behavior*, 114, 106553.
- Stokel-Walker, C. (2023). ChatGPT listed as author on research papers: Many scientists disapprove. *Nature*, 613(7945), 620-621.
- 牟怡, 夏凯, NovoZhilova, E., et al. (2019). 人工智能创作内容的信息加工与态度认知: 基于信息双重加工理论的实验研究. *新闻大学*, (8), 30-43, 121-122.
- 方师师, 唐巧盈. (2023). 聪明反被聪明误: ChatGPT 错误内容生成的类型学分析. *新闻与写作*, (4), 31-42.
- 朱君辉, 王梦焰, 杨尔弘, 等. (2024). 大模型生成回答与人类回答文本的语言特征比较研究. *中文信息学报*, 38(4), 17-27.
- 孙利萍, 方清明. 汉语话语标记的类型及功能研究综观. *汉语学习*, 2011(6), 76-84.