Journal of Contemporary Educational Research, 2023, Volume 7, Issue 12
http://ojs.bbwpublisher.com/index.php/JCER
ISSN Online: 2208-8474
ISSN Print: 2208-8466

Corpus-Driven Analysis of Conceptual Metaphor in Artificial Intelligence Language: A Sample of ChatGPT-Written Speeches

Yang Yang*

School of Foreign Languages, Qinghai University, Xining 810016, Qinghai Province, China

***Corresponding author:** Yang Yang, cstdxy2008@126.com

Abstract: Based on Conceptual Metaphor Theory (CMT), this paper creates a tiny corpus of ChatGPT-written speeches. Through employing a corpus-driven approach, this study analyzes the identification and utilization of conceptual metaphors in artificial intelligence (AI) languages. The AI demonstrated its capacity to utilize metaphors in the metaphoric corpora through the display of diversity, non-arbitrariness, repetition, and intersectionality in the selection of source domains. It often uses vocabulary combinations with clear similarities to establish metaphorical meaning. In the literal sense, the outcomes of metaphor identification by artificial intelligence differ significantly from those of humans. Therefore, there is a need to develop advanced automatic models for identifying metaphors in order to enhance the precision of metaphor identification consistently.

Keywords: Artificial intelligence language; ChatGPT; Conceptual metaphor; Identification

Online publication: December 25, 2023

1. Introduction

OpenAI's Artificial Intelligence Laboratory introduced a chatbot program known as ChatGPT (Chat Generative Pre-Trained Transformer) on November 30, 2022. This application boasts a wide range of functionalities, including text generation, code generation, and language translation, all driven by user input. In just a short span of two months since its release, ChatGPT has already amassed over 100 million active users per month. This remarkable achievement positions ChatGPT as one of the fastest-growing consumer-level applications in history [1]. Moreover, its user base extends across diverse sectors and fields within society, exerting a far-reaching influence and significantly shaping perceptions of the effectiveness of cutting-edge artificial intelligence (AI) technologies on a global scale. It has ignited a technological revolution in the field of AI [2]. The evolution of artificial intelligence has ushered in a new era.

Despite the widespread user experience assessments conducted by the public, the author has observed a noticeable dearth of research on ChatGPT within the academic field. While evaluations within the computer-

related technology sector abound, there is a significant scarcity of research in the humanities and social sciences. Specifically, there is a lack of readily accessible research on artificial intelligence language from a linguistic standpoint. As artificial intelligence advances and its applications become increasingly prevalent, the language associated with artificial intelligence becomes more pervasive in society, extending its influence on language. Therefore, a comprehensive investigation of the language utilized in artificial intelligence warrants considerable focus.

The primary focus of ChatGPT's interface is centered on language and textual content. It functions as genuine communications resulting from contact between humans and machines, and as a crucial component of contemporary society communication. Therefore, it can be considered a fundamental study topic in the field of artificial intelligence language. This paper employs the theoretical framework of conceptual metaphors and examines an English speech generated by ChatGPT as a case study. The objective is to examine matters on the identification and utilization of metaphors in artificial intelligence. This involves referencing data to unveil the cognitive representations in artificial intelligence language and improve models for metaphor identification.

2. Conceptual Metaphor Theory

The English word "metaphor" originated from the Greek word "metaphora," which is a compound of "meta" and "pherein." In this compound, "meta" means "over," and "pherein" means "to carry." Therefore, the original meaning of "metaphor" suggests a form of "movement from here to there," symbolizing a type of transformation [3]. In traditional theories, three primary schools of thought address the origin of metaphors: the comparison view, the substitution view, and the interaction view. A notable departure from traditional metaphor theories is the cognitive perspective on metaphors, known as Conceptual Metaphor Theory. This theory originated from the collaborative work of Lakoff and Johnson in their book *Metaphors We Live By*. According to Kovecses, the dominance of Conceptual Metaphor Theory can be attributed to its close alignment with "a variety of disciplines and approaches in the study of the human mind and human behavior" [4]. Lakoff posits that "metaphor permeates every facet of human existence, manifesting not only in our language but also in the core of our thoughts and behaviors. Fundamentally, the conceptual systems that underpin human cognition and conduct are metaphorical in nature" [5]. According to CMT, its main viewpoints include:

(1) Metaphors are all-pervasive

(2) Systematic mappings between two conceptual domains

(3) From concrete domain to abstract domain

(4) Metaphors occur primarily in thought

(5) Conceptual metaphors are grounded

(6) Provenance of source domains [6]

Sun Yi posits that a metaphor involves a systematic mapping from a well-known, easily understood source domain to a less familiar and less readily comprehensible target domain. Between the source domain and target domain, there exists a series of ontological or cognitive correspondences, often referred to as "mappings." For example, in the metaphor of "love is a journey," the word "journey" is used as the source domain for understanding and describing the target domain of "love" through correspondence and mapping. The explanatory power of the Conceptual Metaphor Theory is convincing, and it possesses the advantages of efficiency and conciseness in analyzing and unearthing metaphorical meanings and entrenched concepts [7].

Over the past four decades, Conceptual Metaphor Theory has thrived, growing in terms of its disciplinary content (including research issues, perspectives, and methods) and societal applications [8]. The influence of

Conceptual Metaphor Theory has been substantial across multiple disciplines, including linguistics, psychology, philosophy, and cognitive science. This text emphasizes the significance of metaphor in everyday language, cognition, and communication, and its impact on our comprehension of abstract ideas.

3. Corpus collection

In order to uphold the scientific rigor and precision of this study, the author initially posed two queries to ChatGPT:

(1) What does cognitive linguistics' Conceptual Metaphor Theory entail?

(2) Is it possible to employ or abstain from employing metaphors according to this theory?

Upon receiving accurate and positive responses from the AI, the research advanced to the subsequent stages.

The author directed ChatGPT to generate English addresses utilizing Conceptual Metaphor Theory, encompassing ten subjects: education, ecology, poverty, culture, gender, science, corruption, economy, health, and politics. ChatGPT was assigned the task of writing two groups of speeches for each topic—one utilizing metaphors and the other without. Consequently, there was a total of 20 speeches. Following that, the author eliminated the opening and closing parts of the speeches, such as "Ladies and gentlemen" and "Thank you." This culminated in a limited collection of English speeches created by ChatGPT, containing 10,553 words and 10,599 tokens (**Table 1**).

Table 1. Token and types in two groups of English speeches for ten topics

Group Topics	Group 1 Without metaphor usage		Group 2 With metaphor usage	
	Token	Type	Token	Type
Education	438	437	436	438
Ecology	477	480	398	400
Poverty	418	420	532	533
Culture	593	595	491	493
Gender	540	543	547	550
Technology	620	623	644	646
Corruption	609	611	436	436
Economy	537	538	530	532
Health	502	509	630	635
Politics	643	645	532	535
Total	5377	5401	5176	5198

4. Steps of metaphor identification

During the metaphor identification phase, the author applied the "Metaphor Identification Procedure of VU University Amsterdam (MIPVU)" [9], a widely acknowledged and extensively utilized method, to assure the research's standardization and comprehensiveness. The specific procedures were as follows: Firstly, the discourse text was read word by word to determine the vocabulary that requires annotation; Secondly, when a word was used indirectly and this usage can be explained through cross-domain mapping from a more basic

sense, the term was annotated as “indirect metaphor”; Thirdly, when a word was used directly and this usage can be explained through cross-domain mapping from a more basic sense, the term was annotated as “direct metaphor”; Fourthly, when vocabulary-grammatical substitutions were involved, and the direct or indirect meaning conveyed by the replaced or omitted vocabulary can be explained through cross-domain mapping from a more basic sense, result, or topic, the term was annotated as “implicit metaphor.” The basic sense here tends to refer to the MIP’s basic steps and exhibits the following characteristics: greater specificity, relevance to bodily actions, higher precision, and a historical occurrence.

Following the completion of thorough identification of metaphors in the corpus, an assessment of inter-rater reliability was carried out. The present evaluation utilized Pearson’s correlation analysis, a widely recognized statistical technique commonly applied in research endeavors that involve multiple raters. The objective of this analysis was to determine the degree of agreement among the raters. The strong degree of concurrence noted among the evaluators in this examination provides empirical support for the dependability of the identified metaphors. Subjective bias may manifest in individual metaphor identifiers during the identification procedure.

5. Metaphor analysis in artificial intelligence language

The author provided ChatGPT with two groups of speech for each subject matter: one containing metaphors and the other devoid of them. The purpose of this analysis was to examine the attributes of metaphor usage and the accuracy of metaphor identification in artificial intelligence.

5.1. Instructions for ChatGPT to use metaphors in speeches

By undertaking manual identification and verification, the author discovered that metaphorical words occurred 628 times in total across the ten topics, constituting 12% of the text. The frequencies of metaphorical words in each topic, ranked from highest to lowest, are as follows: economy (89), health (78), education (75), politics (66), corruption (65), technology (58), gender (54), culture (53), poverty (53), and ecology (37), as displayed in **Table 2**. ChatGPT employed a greater number of metaphorical phrases when discussing economic subjects while using fewer metaphorical words when discussing ecological subjects.

Table 2. Frequency and token of metaphorical words in ten topics

Topics	Frequency of metaphorical words (times/ occurrences)	Token (times/ occurrences)	%
Education	75	436	17%
Ecology	37	398	9%
Poverty	53	532	10%
Culture	53	491	11%
Gender	54	547	10%
Technology	58	644	9%
Corruption	65	436	15%
Economy	89	530	17%
Health	78	630	12%
Politics	66	532	12%
Total	628	5176	12%

It is worth mentioning that according to the Conceptual Metaphor Theory, metaphors involve two cognitive domains, and their essence consists of the systematic cross-domain mappings that move from a source domain to a target domain [10], which needs to be taken into consideration. Due to this, one of the primary focuses of metaphor research is an investigation of the source domain. The ChatGPT writing corpus was subjected to manual examination and classification, and the results revealed that it contains a total of forty unique source domains, each of which relates to a different human experience that is intimately connected to human life. The diverse source domains that are covered include planting, travel, architecture, containers, space, commerce, tools, the human body, time, power, courses, textiles, drama, systems, war, accidents, traffic, religion, gifts, sports, demons, illnesses, disasters, light, color, sound, food, fine arts, astronomy, literature, animals, clothing, environmental protection, companionship, games, professions, hazardous substances, and machinery.

In addition, statistical analysis of the study of the source domain usage frequency in ChatGPT discovered that the usage frequencies of the various source domains differ greatly from one another. The following is a breakdown of the specific statistics: In total, 8 source domains have usage frequencies of 8 times, they are business (9), planting (9), architecture (9), space (9), war (9), trips (8), the human body (8), and power (8). Moreover, 6 source domains have usage frequencies ranging from 4 to 7 times (inclusive), including textiles (7), diseases (6), natural phenomena (6), drama (5), tools (5), and color (4). Meanwhile, containers (3), time (3), courses (3), systems (3), sports (3), music (2), animals (2), and another 26 source domains all have usage frequencies of at least three times (**Figure 1**). According to the statistics, it is clear that source domains such as commerce, planting, architecture, space, and war exist in a variety of themes with high usage frequencies. This indicates that ChatGPT's responses demonstrate a recurrent exploitation of source domains.

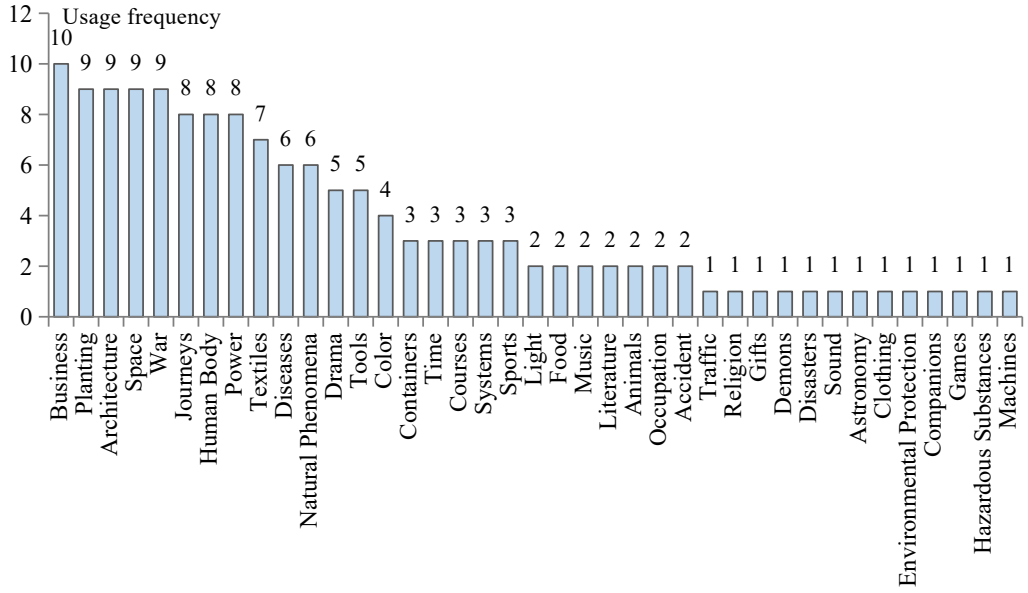

Figure 1. Frequency statistics of 40 source domains

A more thorough analysis of the utilization rates of the 40 source domains revealed that the number of source domains employed in different themes varies. In the political domain, 21 source domains were used, constituting 52% of the total. In the education and health domains, 11 source domains were engaged in each, amounting to 27.5% each (**Table 3**).

ChatGPT exhibits a greater propensity to select source domains for relatively abstract and less readily comprehensible subjects. Conversely, for more specific and readily comprehensible topics, the availability of

source domains is comparatively constrained. This demonstrates a non-arbitrary selection pattern. Based on this observation, it can be inferred that the quantity of source domains chosen by ChatGPT reflects its level of understanding of the target domain. If a greater number of source domains are selected, it indicates that ChatGPT finds the target domain more challenging to comprehend, and vice versa.

Table 3. Number of source domains in ten topics

Topics	Number of source domains
Education	11
Ecology	13
Poverty	13
Culture	14
Gender	15
Technology	13
Corruption	20
Economy	15
Health	11
Politics	21
Total	147

Additionally, within ChatGPT's written corpus, there are instances where multiple source domains establish similar relationships with a single target domain and engage in cross-domain mapping. Consequently, the questions arise: Do different source domains within the same topic have similar usage frequencies? Which source domains are the most representative? What relationships exist between typical source domains? These questions are worth investigating.

Resonance is a crucial criterion for measuring the frequency of a specific metaphor in a given corpus [11]. Higher resonance values indicate higher productivity of the metaphor, signifying a greater coverage of that metaphor in the specific corpus. If we consider the metaphor-carrying words within a particular source domain type as "types," and their frequencies of occurrence in various forms as "tokens," with Σ representing summation, the formula for calculating "Source Domain Resonance" is as follows: Resonance = Σ(type) × Σ(token). Using this method, the resonance values for all source domains involved in each topic were calculated to identify the most representative source domains in the ten topics (**Table 4**).

Table 4. Source domains with the highest metaphor productivity in ten topics

Topics	Source domain	Token	Type	Resonance
Education	Planting	28	19	532
Ecology	Disease	11	10	110
Poverty	Planting	9	9	81
Culture	Textiles	13	7	91
Gender	Textiles	21	11	231
Technology	Textiles	10	7	70
Corruption	Planting	12	8	96
Economy	Planting	47	27	1269
Health	Planting	48	26	1248
Politics	Music	21	13	273

It is clearly shown in **Table 4** that the "planting" source domain has the highest metaphor productivity in education, poverty, corruption, economy, and health topics. Within the topics of culture, gender, and technology, the source domain of "textiles" exhibits the greatest level of metaphorical productivity. Within the field of ecology, the source domain of "disease" exhibits the highest level of metaphorical productivity. Lastly, in politics, the source domain of "music" has the highest level of metaphorical productivity. According to this, it can be inferred that the "planting" source domain is common in the fields of education, poverty, corruption, economy, health, and other related topics. The domain of "textiles" is often regarded as a prototypical source domain in the fields of culture, gender, and technology. The primary source domain in the field of ecology is the concept of "disease." The presence of these common source domains spans various subjects, indicating the recurring and overlapping patterns in ChatGPT's selection of typical source domains.

Moreover, by examining the concordance of the corpus, the author discovered that ChatGPT consistently used terms such as "as," "like," and "is" to create connections across different domains, resulting in the creation of metaphorical meanings. The most prevalent form of generating metaphorical connotations is the utilization of "like" (100%), as presented in **Table 5**.

Table 5. Statistics of common metaphorical word collocation

Words	Occurrences	%
as	7	18%
is	80	65%
like	54	100%

In summary, ChatGPT has demonstrated the ability to select a diverse, non-arbitrary, repeated, and overlapping variety of source domains based on human experiences closely related to human life in tasks that include metaphorical writing through continuous learning and training on large datasets. In the establishment of metaphorical meanings, ChatGPT frequently uses word combinations that exhibit clear connotations of analogy, which align with common expressions found in human metaphorical language. This leads to the inference that artificial intelligence has acquired the capability for metaphor usage and demonstrated a cognitive processing approach similar to human interaction, wherein humans utilize their bodily experiences in their environment for cognitive processing.

5.2. Analysis of ChatGPT's responses without metaphors

It is noticeable that despite the instruction to generate ten speeches without using metaphors, there were instances of metaphorical terms in the generated text. The cumulative frequency of metaphorical words utilized reached 337, accounting for 6.27% of the text. Out of all the corpora, the one focused on corruption used the most metaphorical words, with a total of 61, making up 10.02% of the corpus (**Table 6**).

Table 6. Results of manual metaphor identification (responses without metaphor usage)

Topics	Tokens of metaphorical words (occurrences)	Metaphorical word usage frequency
Education	23	25.5%
Ecology	15	3.14%
Poverty	26	6.22%
Culture	24	4.05%

Table 6. (Continued)

Topics	Tokens of metaphorical words (occurrences)	Metaphorical word usage frequency
Gender	33	6.11%
Technology	38	6.13%
Corruption	61	10.02%
Economy	36	6.70%
Health	29	5.78%
Politics	52	8.09%
Total	337	6.27%

Based on the results of the independent samples *t*-test, a *P* value of 0.00000045619 was obtained, indicating that the results have statistical significance as the *P* value is less than 0.05. Consequently, it can be inferred that there is a significant difference in metaphor recognition between artificial intelligence and manual recognition, with a substantial disparity. The reasons for this discrepancy may be attributed to the following three factors:

(1) Programming flaws: In highly specialized and academic fields, artificial intelligence may not write according to the specific instructions provided.

(2) Comprehension bias: Although artificial intelligence can respond to academic concepts and theoretical content, it may not deeply comprehend their essence and apply theories to practical contexts.

(3) Rigid patterns and methods of artificial intelligence recognition: Artificial intelligence may employ traditional rhetorical techniques such as association and metaphor but may not follow the structured steps of manual recognition programs.

These factors can contribute to issues of metaphor recognition bias and low recognition accuracy in artificial intelligence. Hence, it is essential to improve and refine the understanding and recognition of metaphors in artificial intelligence, constructing high-performance automatic metaphor recognition models, enhancing the reliability of metaphor labeling, and reducing the burden of manual labeling in natural language processing.

6. Conclusion

This study investigated the utilization and identification of metaphors in the ChatGPT speech corpus using corpus methods, yielding valuable data to enhance the cognitive representation of metaphors and models for automatic metaphor identification in artificial intelligence language. However, this research has limitations, including constraints related to time and human resources. Future studies may explore the construction of diverse large-scale artificial intelligence language corpora and validate hypotheses through a combination of introspection, corpus methodologies, and computational simulations. The examination of artificial intelligence language is not solely the domain of computer science scholars; it also carries a societal responsibility within the field of linguistics. Linguists should focus not only on the fundamental aspects of artificial intelligence language but also delve into language ethics, ideologies, aesthetics, and related matters in response to this emerging field. Addressing new challenges and issues in research requires a comprehensive investigation across multiple dimensions and innovative approaches that bridge different disciplines.

Disclosure statement

The author declares no conflict of interest.

References

[1] Zhang W, Liu T, 2023, Technical Analysis of ChatGPT and Perspectives on the Development of Artificial General Intelligence. Bulletin of National Science Foundation of China, 37(5): 747–753.

[2] George AS, George AH, 2023, A Review of ChatGPT AI's Impact on Several Business Sectors. Partners Universal International Innovation Journal, 1(1): 9–23.

[3] Shu D, 2000, Research on Metaphor, Shanghai Foreign Language Education Press, Shanghai.

[4] Kovecses Z, 2018, Metaphor in Media Language and Cognition: A Perspective from Conceptual Metaphor Theory. Lege Artis, 3(1): 124–141.

[5] Lakoff G, Johnson M, 1980, Metaphors We Live By, Chicago University Press, Chicago.

[6] Kovecses Z, 2016, Conceptual Metaphor Theory, in The Routledge Handbook of Metaphor and Language, Routledge, London, 13–27.

[7] Sun Y, 2019, Constructing Theoretical Paradigms of Contemporary Metaphorology, Hainan: Humanities & Social Sciences Journal of Hainan University, 37(6): 126–134.

[8] Zhang W, 2023, Research on Metaphor and Metonymy, Foreign Language Teaching and Research Press, Beijing, 5.

[9] Steen GJ, 2010, A Method for Linguistic Metaphor Identification, John Benjamins Publishing Company, Amsterdam.

[10] Lakoff G, 1993, The Contemporary Theory of Metaphor, Metaphor and Thought, Cambridge University Press, Cambridge, 202.

[11] Charteris-Black J, 2004, Corpus Approaches to Critical Metaphor Analysis, Palgrave Macmillan, New York, 89.