Figar, V. (2026). Metaphorical framings in newspaper articles generated by ChatGPT-5.6 and Gemini, dealing with the war in Iran, *SCIENCE International journal, 5*(3), 349-355.
doi: 10.35120/sciencej0503349f UDK: 070:355.4(55)]:004.8:004.775

METAPHORICAL FRAMINGS IN NEWSPAPER ARTICLES GENERATED BY CHATGPT-5.6 AND GEMINI, DEALING WITH THE WAR IN IRAN

Vladimir Figar[1*]

[1]English Department, Faculty of Philosophy, University of Niš, Serbia, e-mail: vladimir.figar@filfak.ni.ac.rs

Abstract: The paper aims to explore the range of metaphorical framings in newspapers articles generated by ChatGPT-5.6 and Gemini, dealing with the war in Iran. We have compiled a small specialized corpus with 42 articles and the total corpus length of 19,758 words. LLMs were given topically related article titles and introductory paragraphs chosen randomly from various online sources (e.g., Al Jazeera, Reuters). Based on the supplied information, the LLMs were instructed to generate three different versions of the article, from the following three viewpoints: (i) neutral, (ii) pro-Iranian, and (iii) pro-Israeli. All articles were tagged manually for instances of metaphorically used words, and prepared for subsequent analyses using WordSmith Tools 6.0. Metaphor identification was conducted using the MIPVU (Steen et al., 2010). Quantitative analysis showed a significant main effect of LLMs in the neutral condition (p=.01), with a significantly higher overall mean density of metaphorically used words in articles generated by Gemini. The difference in metaphor density in the two biased-framing conditions did not reach significance. Qualitative analysis revealed differences in the range and types of metaphorical framings between the two LLMs, where the discourse generated by Gemini showed more affective components compared to articles generated by ChatGPT-5.6. The discourse generated by ChatGPT-5.6 was dominated by CONTAINMENT, FORCE, and JOURNEY metaphors. While these metaphor groups were also quite frequent in the discourse generated by Gemini, the latter LLM also generated metaphorical expressions corresponding to FIRE, HYDRA, CANCER, DISEASE, and SNAKE metaphors. Through training and fine-grained instructions, the researched LLMs can pose as useful tools for researching possible framings and their rhetorical functions in discourse.

Keywords: *metaphorical expression, framing, MIPVU, LLM discourse.*
Field: Humanities

1. INTRODUCTION

Conceptual metaphor has provided fruitful ground for investigations in the domains of cognitive linguistics, psycholinguistics and discourse studies. Evolving from a mere linguistic ornament, metaphor established itself as one of the key cognitive mechanisms that facilitate the construction of meaning and everyday interaction (Lakoff & Johnson, 2003[1980]). The range of metaphorical framings generated by LLMs should provide valuable insight into the manner in which LLMs process and generate metaphor. The present paper aims to explore the range of metaphorical framings generated by ChatGPT-5.6 and Gemini in a specific context of newspaper articles dealing with the war in Iran. This topic can be approached from multiple perspectives which could yield different metaphorical framings. The theoretical framework of the paper includes conceptual metaphor theory (Lakoff & Johnson, 2003[1980]), discourse approaches to critical metaphor analysis (Charteris-Black, 2004, 2011), and frame semantics (Fillmore, 1982).

Conceptual metaphor is grounded in the notion of unidirectional, partial, systematic cross-domain mappings from the source domain, which is typically more tangible (e.g., JOURNEY), to the target domain (e.g., LIFE), which is more abstract (Lakoff & Johnson, 2003[1980]). An important distinction is made between a conceptual metaphor which poses as a cognitive mechanism, and a metaphorical expression, which represents concrete linguistic realizations of the higher-order mappings (Lakoff & Johnson, 2003[1980]). Based on the degree of entrenchment a distinction is also made between novel and conventional metaphors, where the latter have become conventionalize through repeated use (Charteris-Black, 2004, p. 17). The initially introduced sharp distinction between structural, orientational and ontological metaphors was later abandoned, as Lakoff & Johnson (2003[1980], p. 264) concluded that all metaphors share characteristic of each of the three groups. Charteris-Black (2004, p. 244) introduced a hierarchy of metaphorical mappings with conceptual keys as the most inclusive (e.g., POLITICS IS CONFLICT), conceptual metaphors, which are more specific, but still pose as cognitive mechanisms (e.g., POLITICS IS A BATTLE), and metaphorical expressions.

The development of frame semantics can be roughly traced through the following stages: (i) the scene and frame model, (ii) the connection between framing and categorization, (iii) the most widely quoted version of the theory from Filmore (1982), and (iv) the relationship between the semantics of

*Corresponding author: vladimir.figar@filfak.ni.ac.rs

understanding and the semantics of truth (Figar, 2021, pp. 85–92). Semantic frame is defined as "any system of concepts related in such a way that to understand any one of them you have to understand the whole structure in which it fits; when one of the things in such a structure is introduced into a text, or into a conversation, all of the others are automatically made available" (Fillmore 1982: 111). The approach also incorporates the import of situational context in the process of meaning construction. Additionally, framing is closely related to the notions of perspective and viewpoint, as the process of meaning construction can typically take place from multiple viewpoints. In that sense, Filmore (1982, p. 124) distinguishes between invoked frames, imposed by the writer or narrator, and evoked frames, related to the relevant knowledge structures that the reader might decide to activate in the process of meaning construction. Figar (2020) conducted a study that involved a response time paradigm with semantic priming and an online categorization task where he provided evidence in favor of the psychological reality of semantic frames. The concise overview of the main tenets of the theoretical framework will be followed by the overview of some of the relevant studies in the field of discourse analysis and LLMs.

Ichien, Stamenković, & Holyoak (2024) tested the ability of GPT-4 to interpret metaphors from novel English poems, and metaphors translated into English from Serbian poetry. The assessment was based on the Fig-QA dataset. The results showed that GPT-4 performed better than previous LLMs. Moreover, human judges assessed metaphor interpretation provided by GPT-4 as better compared to the interpretations obtained from college students. In the case of novel English poems, the obtained interpretations were assessed as good or excellent. In an earlier study, Liu et al. (2022, p. 4437) used the same method to explore the ability of LLMs to interpret figurative language, and they found that "although language models achieve performance significantly over chance, they still fall short of human performance, particularly in zero- or few-shot settings." Muñoz-Ortiz, Gómez-Rodríguez, & Vilares (2024) compared news texts written by humans to those generated by six various LLMs, and the results showed that sentence lengths in texts written by humans ere more uneven, humans used a greater variety of vocabulary items, shorter constituents, dependency distances were more optimized. On the other hand, LLM generated texts contained more symbols, numbers, auxiliaries and pronouns compared to texts written by humans. Dönmez et al. (2025) analyzed how LLMs organize their arguments in persuasive contexts, with the goal of providing a framework for distinguishing human written texts from LLM generated texts. Specifically, the authors analyzed general-purpose linguistic features and domain specific features. The former included elements like syntactic complexity, while the later included engagement strategies etc. They analyzed texts written by humans and texts generated by three different LLMs, and the results showed that "LLM-generated counter-arguments show lower type-token and lemma-token ratios but higher emotional intensity – particularly in anticipation and trust" (Dönmez et al., 2025, p. 34595). Additionally, the argumentation provided by the LLMs was of a textbook kind, and the provided counterarguments seemed to mimic the style of writing to which they were replying. The authors propose the identified differences can be used to detect LLM generated writing. Finally, Breazu & Katsos (2024) compared ChatGPT-4 generated newspaper articles dealing with immigration with actual human written articles. The results showed "that ChatGPT-4 exhibits a remarkable degree of objectivity in its reporting and demonstrates heightened racial awareness in the content it produces" (Breazu & Katsos, 2024, p. 687). Moreover, the generated texts were based on objective, factual information, rather than sensationalism.

With the overview of the theoretical framework and previous research completed, we now turn to methodology and corpus analysis.

2. MATERIALS AND METHODS

We designed a small specialized corpus (in line with Koester, 2010) that included 42 newspaper articles generated by the two LLMs, with a total of 19,758 words. Each LLM was given seven article titles and an introductory paragraph on the topic of the ongoing war in Iran, randomly selected from online news outlets Al Jazeera and Reuters, and each LLM was instructed to generate three different versions of the article, and to utilize metaphorical framings and metaphorical expressions of their own choice that are common in newspaper reporting. They were also instructed to use newspaper writing style and to generate authentic texts that are not copied from existing sources. Each article version was supposed to contain around 500 words. In effect, we obtained 21 articles from each of the two LLMs, with the overall mean article length of 470.23 words (SD=43.27).

In the following step, we analyzed all articles for instances of metaphorically used words, and tagged them manually and prepared them for subsequent analyses using WordSmith Tools 6.0 (Scott, 2010). Metaphor identification was conducted in line with the MIPVU procedure (Steen et al., 2010, pp. 25–42), which entailed the following: (i) reading the entire article in order to establish the relevant context,

(ii) identification of lexical items, (iii) comparison of contextual and basic meanings of target items in order to determine whether the target item has been used in the metaphorical sense, i.e., whether its contextual meaning could be viewed as a function of its basic meaning, and (iv) the analysis was conducted in two passes, with fourteen days between them. In line with Figar (2021, p. 2011), we employed "the provisional operational annotation of the identified metaphors as members of specific groups of conceptual metaphors, and their overarching conceptual keys." We are aware that the identification of conceptual mappings represents a separate research question (in line with Steen et al., 2010), and our provisional annotation should be understood as an artifact of analysis, not an attempt to argue in favor of the existence of conceptual mappings.

The adopted approach was expected to provide answers to the following research questions: (i) are there any differences in the density of metaphorical expressions in articles generated by the two LLMs in conditions of neutral and biased framings?, and (ii) are there any differences in the range of metaphorical expressions generated by the two LLMs in terms of the main conceptual keys, and are there any differences in their rhetorical functions?

3. RESULTS

Descriptive quantitative analysis conducted using the concordance and search-over-tags in WordSmith Tools 6.0 showed the overall mean density of 171.29 (SD=47.08) metaphorically used words per 1,000 words for the entire corpus. One-way ANOVA did not reveal any significant differences in the overall mean metaphor densities between articles generated by the two LLMs (MGPT=165.23, SDGPT=48.79, MGEMINI=177.36, SDGEMINI=45.68, p=.41). One-way repeated measures ANOVA showed a significant effect of LLMs only for the neutral condition, with a significantly higher overall mean metaphor density recorded in articles generated by Gemini (MGPT=156.97, SDGPT=49.60, MGEMINI=194.58, SDGEMINI=41.60, p=.01), while the other two framing conditions did not produce significant effects in terms of metaphor density between the two LLMs (pPRO-IRANIAN=.32, pPRO-ISRAELI=.46).

The comparison of the range of metaphorical framings utilized by the two LLMs showed that ChatGPT-5.6 generated 68 different metaphor groups, while Gemini used 92 different metaphor groups. Metaphor groups with the highest overall density for both LLMs include the following: FORCE, CONTAINMENT, JOURNEY, STRUCTURE, CONFLICT, PERSONIFICATION, TRADE, MACHINE, EMISSION, and SPACE. It can be concluded that these are actually some of the common metaphorical framings found in human written newspaper discourse (based on Charteris-Black, 2004). In turn, the conceptual base of metaphorical framings generated by the LLMs appears to mirror the structure of typical newspaper discourse, which is also evident from examples 1–6 below. Namely, peace talks are conceptualized as a moving object (examples 1 and 5), time is conceptualized as a weapon (example 2), economy is conceptualized as a human being threatened by strangulation (example 3), Tehran is conceptualized as an injured person (example 4), arguments are conceptualized as weapons (example 6), and negotiations are conceptualized as boxing (example 3). All of the highlighted metaphorical expressions can be understood as instances of conventional metaphors which are quite common in the political discourse of daily newspapers (Figar, 2021; Charteris-Black, 2004).

1. For now, the talks appear to be advancing<m-JOURNEY>. (ChatGPT-5.6)

2. Washington and Israel are effectively trying to turn<m-FORCE> time into a weapon<m-CONFLICT>. (ChatGPT-5.6)

3. Yet previous rounds<m-boxing> of maximum pressure<m- FORCE > offer little evidence that economic strangulation<m-CONFLICT><m-PERSONIFICATION> automatically produces political surrender<m-CONFLICT>. (ChatGPT-5.6)

4. Tehran is currently bleeding<m-PERSONIFICATION> from a massive internal<m-CONTAINMENT> wound. (Gemini)

5. The Biden-Trump transition team and its Israeli allies are approaching<m-JOURNEY> these Pakistan-mediated talks with deep, battle-tested<m-CONFLICT> skepticism. (Gemini)

6. The escalating rhetorical crossfire<m-CONFLICT> presents a […] logistical dilemma. (Gemini)

In the neutral framing condition, the following metaphor groups appeared in texts generated by both LLMs: CONFLICT, CONTAINMENT, JOURNEY, EMISSION, SPACE, SPORTS, FORCE, WEIGHT, BALANCE, CENTER-PERIPHERY, VERTICALITY, SCALE, STRUCTURE, COMPOSITIONALITY, EXPERIMENT, FIRE, GAMBLE, LIGHT, MACHINE, MATH, NARRATIVE, PERSONIFICATION, VISION, BODY, TEMPERATURE, THEATER, TRADE, VIEWPOINT, WEATHER, and LIQUID. Such findings suggest that both LLMs relied on some of the most common metaphorical framings typically found in

political newspaper discourse, which include CONFLICT, JOURNEY, CONTAINMENT and FORCE (Charteris-Black, 2004). Metaphorical framings identified only in articles generated by ChatGPT-5.6 included: WARNING, TIME, TOOL, SHADOW, WARMTH, NEGOTIATION, SMOKE, DECORATION, CONDUIT, CONTEST, TIME, SOFT, and TEST, while the framings that occurred only in articles generated by Gemini included: ANIMAL, CARROT-STICK, DANCE, CONSTRUCTION, BREATHING, FORTRESS, BAROMETER, WINDOW, DISSECTION, GAME, HAWK, MEASUREMENT, PLANT, LANDSCAPE, HUB, MEDICINE, RESERVOIR, POSTURE, WAR, and WAVE. This overview shows a wider range of metaphorical framings generated by Gemini compared to ChatGPT-5.6 which seems to have been anchored in a more restrained journalistic tone, which is in line with Breazu & Katsos (2024). Also, the framings that occurred only in neutral articles generated by ChatGPT-5.6 are not as emotionally charged as those that were exclusive to articles generated by Gemini. While metaphorical framings presented in examples 7–10 were not very frequent, conceptualizing the political entities as hawks (example 7), the negotiation process as a dissection (example 8), the case of sticks and carrots (example 9), or as a dance (example 10) could potentially serve to direct the readers' attention towards the hostile and unfair aspects of the ongoing crisis. In addition, these idiosyncratic framings have been successfully integrated into the discourse with the more conventional CONTAINMENT, JOURNEY, EMISSION, PERSONIFICATION, and COMPOSITIONALITY metaphors, thereby forming a coherent metaphorical scenario, similar to that identified with metaphor clusters (in the sense of Kimmel, 2010). Namely, the metaphorical framings that are related through the situational context most likely serve to reinforce the rhetorical content of the intended message.

7. The strict benchmarks demanded by congressional hawks<m-HAWK> represent an effort to drive up the cost of entry<m-CONTAINMENT><m-JOURNEY> for Tehran. (Gemini)

8. The mixed signals emanating<m-EMISSION> from<m-CONTAINMENt> both capitals […] are being closely dissected<m-DISSECTION> by foreign intelligence agencies. (Gemini)

9. This […] reflects a classic institutional debate over the optimal mix<m-COMPOSITIONALITY> of diplomatic carrots and structural sticks<m-CARROT-STICK>. (Gemini)

10. This intricate dance<m-DANCE> between Washington and Tehran highlights how deeply interconnected modern security dynamics is with global macroeconomic health<m-PERSONIFICATION>. (Gemini)

The pro-Iranian framing condition showed that the following metaphorical framings occurred with both LLMs: CONFLICT, CONTAINMENT, JOURNEY, FORCE, BALANCE, COMPOSITIONALITY, STRUCTURE, SPACE, VERTICALITY, VIEWPOINT, BODY, ANCHOR, EMISSION, LIGHT, FOOD, TOOL, GAME, HAWK, MACHINE, NARRATIVE, ORIENTATION, VISION, THEATER, PERSONIFICATION, WEATHER, and WAVE. On the other hand, certain metaphorical framings were exclusive to one of the two LLMs, with BOXING, FOOTBALL, GAMBLE, CENTER-PERIPHERY, SCALE, NEGOTIATION, SHADOW, STONE, WARNING, SOFT, TABLE, CLOTHING, and COURT appearing only in articles generated by ChatGPT-5.6. Gemini generated the following unique metaphorical framings: ADULTHOOD, BREATHING, DISEASE, ANATOMY, FORTRESS, DREAM, HIJACK, MARKETING, HUBRIS, GIFT, PREDATOR, MONSTER, PLANT, PIRACY, MEDICINE, POISON, ROGUE, SLEEP, RELIC, SURGERY, TIDE, TEMPERATURE, and WEB. Again, the metaphorical framings generated by Gemini appear to have a higher affective potential, which is evident from the following examples:

11. The Iranian nation has stood as a resilient shield<m-FORCE> against a lawless, predatory<m-PREDATOR> invasion initiated by the Western-Israeli axis<m-FORCE>. (Gemini)

12. The frantic opposition from pro-Israel groups merely highlights their terror of losing a manufactured monster<m-MONSTER>, which they constantly use to justify their own illegal occupations. (Gemini)

13. The financial blockade<m-CONTAINMENT><m-FORCE> acts as a venom<m-POISON> injected< m-FORCE > directly into< m-CONTAINMENT > the veins<m-PERSONIFICATION> of the nation's healthcare. (Gemini)

14. Our ballistic response on Friday night was merely an introductory note<m-NARRATIVE> an explicit warning of what will happen if this state-sponsored piracy<m-PIRACY> continues. (Gemini)

15. For Tehran […] that argument is a demand that Iran dismantle<m-STRUCTURE><m-fORCE> its own sovereignty in exchange for promises from governments that have repeatedly shifted<m-FORCE> the goalposts<m-FOOTBALL>. (ChatGPT-5.6)

16. The economic stakes<m-GAMBLE> further complicate Washington's position<m-JOURNEY>. (ChatGPT-5.6)

17. Each threat from Washington has become another stone<m-STONE> dropped into<m-CONTAINMENT> the global financial pond<m-POND>, sending ripples< m-FORCE ><m-JOURNEY>

through equities, bonds… (ChatGPT-5.6)

In example 11, the enemy is framed as a predator, and this metaphor interacts with two force metaphors. In example 13, the financial blockade is conceptualized as venom, and this framing also interacts with force, containment and personification metaphors, most likely in order to provoke sympathy from the readers. In example 14, the enemy is conceptualized as a pirate, which highlights that actions undertaken by Israel are illegal. In example 12, Iran is conceptualized as a "manufactured monster" by its adversaries, which was probably intended to promote the image that would legitimize the military action against it. Examples 15–17 show metaphorical framings generated by ChatGPT-5.6, where political maneuvers are presented as shifting the goal posts, which highlights the unfair nature of the process (example 15), economy is depicted as a gamble (example 16), and threats are conceptualized as stones that affect the global economy (example 17). The identified framings appear to be working in concert with the remaining metaphorical conceptualizations identified in the texts, thereby strengthening the intended rhetorical message.

Finally, the pro-Israeli framing condition appears to have produced the greatest qualitative differences between the two LLMs. Metaphor groups that occurred with both LLMs included the following: CONFLICT, CONTAINMENT, FORCE, JOURNEY, EMISSION, FIRE, SPACE, STRUCTURE, TEMPERATURE, VERTICALITY, VIEWPOINT, WAVE, WEATHER, NARRATIVE, MACHINE, GAME, HAWK, LIFELINE, LIQUID, MACHINE, LIGHT, and PERSONIFICATION. Metaphorical framings unique to ChatGPT-5.6 were: BOXING, CONTEST, CHAIN, FORTRESS, CENTER-PERIPHERY, NEGOTIATION, ORIENTATION, SOFT, TIME, VISION, MATH, LINE, LANDSCAPE, OXYGEN, and EXPERIMENT, while Gemini generated: BEAST, CANCER, CARROT, DENTISTRY, DISEASE, ACTOR, ANCHOR, HYDRA, FOOD, HEART, GAMBLE, MEDICINE, PARASITE, OCTOPUS, POISON, PLANT, SHELL, SNAKE, STORM, SURGERY, WEIGHT, WALL, MEASUREMENT, ROGUE, SIGN, and WINDOW metaphors. Although not very frequent, the following framings generated by Gemini attest to a systematic, emotionally charged conceptualizations of Iran, most likely aimed at enforcing a specific viewpoint on the readers:

18. The renewed and intensified economic siege<m-CONFLICT><m-CONTAINMENT> is systematically starving<m-PERSONIFICATION> the beast<m-BEAST> of the vital capital it desperately needs to regenerate<m-PERSONIFICATION> its toxic<m-POISON> tentacles<m-OCTOPUS> in Lebanon, Syria, Yemen, and Gaza. (Gemini)

19. The United States [ensured] that any final document forces<m-FORCE> a systematic extraction of Iran's military teeth<m-DENTISTRY>. (Gemini)

20. Iran's regional influence operates like a multi-headed proxy hydra<m-HYDRA>, terrorizing democratic societies. (Gemini)

21. Israel will no longer tolerate hiding behind proxies […] the head of the snake<m-SNAKE> has now been struck directly<m-CONFLICT>. (Gemini)

22. For decades, the radical Iranian regime has operated as a malignant cancer<m-CANCER> metastasizing<m-CANCER> across<m-SPACE> the entirety of the Middle East. (Gemini)

23. Following the horrific, unprecedented wake-up call<m-WAKING> […] the civilized world finally recognized that this virulent infection<m-DISEASE> could no longer be managed<m-FORCE> with naive diplomatic appeasement. (Gemini)

It can be seen that the articles generated by Gemini in the pro-Israeli condition contain far more volatile conceptualizations compared to articles generated by this LLM in the pro-Iranian condition. Namely, Iran is framed as a beast, octopus and poison (example 18), as a hydra (example 20), as a snake (example 21), as a cancer and a virulent infection (examples 22 and 23), and its military is depicted as a tooth that needs to be extracted (example 19). Such a combination of framings generated in articles in the pro-Israeli condition suggests that the language used in these articles is affectively charged and most likely intended to paint a very negative picture of Iran. The most likely goal of such framings is to coax the readers into giving their full support against Iran, while at the same time backgrounding the negative aspects of the Israeli military actions. In other words, such a rhetorical strategy reflects the "highlighting-and-hiding" function of conceptual metaphors outlined by Lakoff & Johnson (2003[1980]). Ultimately, it is up to the readers to decide whether they will accept the invoked framing of the narrative, or evoke their own framing through the activation of the relevant (background) knowledge (in the sense of Fillmore, 1982).

4. DISCUSSION

In this section we provide answers to the main research questions outlined above.

RQ1. Are there any differences in the density of metaphorical expressions in articles generated by the two LLMs in conditions of neutral and biased framings? Quantitative analysis did not reveal any significant differences in the overall metaphor density between articles generated by ChatGPT-5.6 and Gemini (p=.41). The comparison of the two LLMs showed a significantly higher density of metaphorically used words in the neutral condition for articles generated by Gemini compared to articles generated by ChatGPT-5.6 (p=.01). The comparison of the remaining two biased conditions did not show any significant effects of LLMs (pPRO-IRANIAN=.32, pPRO-ISRAELI=.46).

RQ2. Are there any differences in the range of metaphorical expressions generated by the two LLMs in terms of the main metaphor groups, and are there any differences in their rhetorical functions? Qualitative analysis revealed that some of the common conventional metaphorical framings like CONFLICT, JOURNEY, CONTAINMENT, and FORCE were quite frequent in all articles generated by both LLMs. This suggests that the generated conceptualizations reflect those that are typically found in the political discourse of daily newspapers. Combinations of such conventional metaphorical mappings afford the construction of meaning from a particular perspective, i.e., through metaphorical frames invoked by the text. In turn, such framings are intended to produce a specific rhetorical effect. For instance, conceptualizing the political process as a journey could serve to emphasize the development of the negotiations (example 5 above). However, we must note that we take all conclusions presented within the qualitative analysis above with an important caveat. Namely, qualitative analysis can yield only testable hypotheses in terms of the possible rhetorical functions that metaphorical framings might perform. In order to obtained more reliable data, the identified framings should be tested in an experimental setting which would reveal the actual effect they may have on the readers.

Apart from conventional metaphorical framings, each of the LLMs generated some idiosyncratic framings in each of the three conditions. For instance, in the neutral condition, Gemini generated metaphorical conceptualizations that included ANIMAL, CARROT-STICK, DANCE, CONSTRUCTION, BREATHING, FORTRESS, BAROMETER, WINDOW, DISSECTION, GAME, HAWK, MEASUREMENT, PLANT, LANDSCAPE, HUB, MEDICINE, RESERVOIR, POSTURE, WAR, and WAVE, whereas ChatGPT-5.6 generated the following conceptualizations: WARNING, TIME, TOOL, SHADOW, WARMTH, NEGOTIATION, SMOKE, DECORATION, CONDUIT, CONTEST, TIME, SOFT, and TEST. In the pro-Iranian condition, ChatGPT-5.6 generated framings that involved BOXING, FOOTBALL, GAMBLE, CENTER-PERIPHERY, SCALE, NEGOTIATION, SHADOW, STONE, WARNING, SOFT, TABLE, CLOTHING, and COURT; Gemini, on the other hand, generated framings grounded in the following metaphorical conceptualizations: ADULTHOOD, BREATHING, DISEASE, ANATOMY, FORTRESS, DREAM, HIJACK, MARKETING, HUBRIS, GIFT, PREDATOR, MONSTER, PLANT, PIRACY, MEDICINE, POISON, ROGUE, SLEEP, RELIC, SURGERY, TIDE, TEMPERATURE, and WEB. In the final condition, Gemini generated more volatile framings compared to ChatGPT-5.6, and these included the following: BEAST, CANCER, CARROT, DENTISTRY, DISEASE, ACTOR, ANCHOR, HYDRA, FOOD, HEART, GAMBLE, MEDICINE, PARASITE, OCTOPUS, POISON, PLANT, SHELL, SNAKE, STORM, SURGERY, WEIGHT, WALL, MEASUREMENT, ROGUE, SIGN, and WINDOW. ChatGPT-5.6, on the other hand, generated more affectively balanced texts (in line with Breazu & Katsos, 2024) that included the following framings: BOXING, CONTEST, CHAIN, FORTRESS, CENTER-PERIPHERY, NEGOTIATION, ORIENTATION, SOFT, TIME, VISION, MATH, LINE, LANDSCAPE, OXYGEN, and EXPERIMENT.

In terms of the possible rhetorical functions, articles generated by Gemini appear to have a wider range of metaphorical framings, and those framings also seem more emotionally and affectively charged, rendering them more volatile compared to the discourse generated by ChatGPT-5.6. This is especially evident in the pro-Israeli condition, where, among others, Gemini generated framings that included SNAKE, BEAST, HYDRA, CANCER, and DISEASE metaphors. In effect, both LLMs show potential for generating writing ideas along with specific metaphorical framings and critical perspectives that a writer could use as guidelines.

5. CONCLUSIONS

Based on the obtained data, it can be concluded that metaphorical framings generated by the two LLMs mirror to a great extent those identified in previous studies on the political discourse of daily newspapers (e.g., Charteris-Black, 2004; Breazu & Katsos, 2024). In addition to the common conventional conceptualizations, the LLMs also managed to generate some more idiosyncratic framings linked to

specific framing conditions. In that sense, these LLMs could be further explored as potentially useful tools for generating writing ideas, generating texts based on specific instructions, or generating texts with errors that can later be used as teaching materials for writing, grammar and vocabulary in the context of the ELF classroom. Additionally, the two LLMs could also be used as aids in the development of critical thinking by providing argumentation and framing the issue at hand from various viewpoints.

One of the main limitations of the present study is corpus length, so subsequent research should include larger corpora with texts generated by multiple LLMs. This would afford a more thorough insight into the manner in which LLMs generate and incorporate metaphorical framings into texts. Additionally, bearing in mind that qualitative analysis can offer but a set of testable hypotheses concerning the potential rhetorical function of metaphorical framings, future research should also include experimental procedures that would allow the hypotheses to be explored further.

ACKNOWLEDGEMENTS

Prepared as a part of the project “Science, teaching and the labor market on the occasion of the 55th anniversary of the Faculty of Philosophy, University of Niš: English philology and the impact of artificial intelligence,” conducted at the University of Niš – Faculty of Philosophy (No. 373/1-12-01).

REFERENCES

Breazu, P., & Katsos, N. (2024). ChatGPT-4 as a journalist: Whose perspectives is it reproducing? Discourse & Society, 35(6), 687–707. https://doi.org/10.1177/09579265241251479

Charteris-Black, J. (2004). Corpus Approaches to Critical Metaphor Analysis. Palgrave Macmillan.

Donmez, E., Maurer, M., Lapesa, G., & Falenska, A. (2025). AI Argues Differently: Distinct Argumentative and Linguistic Patterns of LLMs in Persuasive Contexts. In Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (pp. 34595–346). Association for Computational Linguistics. https://doi.org/10.18653/v1/2025.emnlp-main.1755

Figar, V. (2020). Testing the Activation of Semantic Frames in a Lexical Decision and a Categorization Task. Facta Universitatis, Series: Linguistics and Literature, 18(2), 159–179. https://doi.org/10.22190/FULL2002159F

Figar, V. (2021). Semantic Frame Activation and Contextual Aptness of Metaphorical Expressions. Unpublished doctoral dissertation. Faculty of Philosophy, University of Niš, Serbia.

Ichien, N., Stamenković, D., & Holyoak, K. (2024). Large Language Model Displays Emergent Ability to Interpret Novel Literary Metaphors. Metaphor and Symbol, 39(4), 296–309. https://doi.org/10.1080/10926488.2024.2380348

Kimmel, M. (2010). Why we mix metaphors (and mix them well): Discourse coherence, conceptual metaphor, and beyond. Journal of Pragmatics, 42(1), 97–115. https://doi.org/10.1016/j.pragma.2009.05.017

Koester, A. (2010). Building small specialized corpora. In A. O’Keeffe and M. McCarthy (Eds.), The Routledge Handbook of Corpus Linguistics (pp. 66–79). Routledge. https://doi.org/10.4324/9780203856949

Lakoff, G., & Johnson, M. (2003[1980]). Metaphors We Live By. The University of Chicago Press.

Liu, E., Cui, C., Zheng, K., & Neubig, G. (2022). Testing the ability of language models to interpret figurative language. In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (pp. 4437–4452). Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.naacl-main.330

Muñoz-Ortiz, A., Gómez-Rodríguez, C., & Vilares, D. (2024). Contrasting Linguistic Patterns in Human and LLM-Generated News Text. Artificial Intelligence Review, 57(265), 1–28. https://doi.org/10.1007/s10462-024-10903-2

Scott, M. (2010). What can corpus software do? In A. O’Keeffe and M. McCarthy (Eds.), The Routledge Handbook of Corpus Linguistics (pp. 136–151). Routledge. https://doi.org/10.4324/9780203856949

Steen, G. J., Dorts, A. G., Herrmann, J. B., Kaal, A. A., Krennmayr, T., & Pasma, T. (2010). A Method for Linguistic Metaphor Identification: From MIP to MIPVU. John Benjamins Publishing Company. https://doi.org/10.1075/celcr.14