BRAIN. Broad Research in Artificial Intelligence and Neuroscience
e-ISSN: 2067-3957 | p-ISSN: 2068-0473
Covered in: Web of Science (ESCI); EBSCO; JERIH PLUS (hkdir.no); IndexCopernicus; Google Scholar; SHERPA/RoMEO; ArticleReach Direct; WorldCat; CrossRef; Peeref; Bridge of Knowledge (mostwiedzy.pl); abcdindex.com; Editage; Ingenta Connect Publication; OALib; scite.ai; Scholar9; Scientific and Technical Information Portal; FID Move; ADVANCED SCIENCES INDEX (European Science Evaluation Centre, neredataltics.org); ivySCI; exaly.com; Journal Selector Tool (letpub.com); Citefactor.org; fatcat!; ZDB catalogue; Catalogue SUDOC (abes.fr); OpenAlex; Wikidata; The ISSN Portal; Socolar; KVK-Volltitel (kit.edu)
2026, Volume 17, Special Issue 1, pages: 328-353
Submitted: February 27th, 2026 | *Accepted for publication:* June 7th, 2026

Cognitive Mechanisms of Contrast Recognition and Interpretation and Their Functional Analogues in Artificial Intelligence Systems

Liudmyla Volodymyrivna Shytyk
Department of Ukrainian Linguistics and Applied Linguistics, Bohdan Khmelnytsky National University of Cherkasy, Cherkasy, Ukraine.
l_shytyk@ukr.net,
https://orcid.org/0000-0001-5941-672X

Liudmyla Petrivna Yuldasheva
Department of Ukrainian Linguistics and Applied Linguistics, Bohdan Khmelnytsky National University of Cherkasy, Cherkasy, Ukraine.
1alimduyl@gmail.com,
https://orcid.org/0000-0002-6561-8827

Anzhelika Kostiantynivna Dosenko
Department of Journalism, V. I. Vernadsky Taurida National University, Kyiv, Ukraine.
dosenko.anzhelika@tnu.edu.ua,
https://orcid.org/0000-0002-5415-1299

Oleksii Valeriiovych Sytnyk
Institute of Journalism, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine.
sytnyk@knu.ua,
https://orcid.org/0000-0002-0853-1442

Myroslava Mykhailivna Malysh
Department of Journalism, V. I. Vernadsky Taurida National University, Kyiv, Ukraine.
malish.myroslava@tnu.edu.ua,
https://orcid.org/0000-0002-6987-5118

Artem Oleksandrovych Halych
Doctor of Philological Sciences
Professor Acting Dean of the Faculty of Social and Humanitarian Sciences State Institution "Luhansk Taras Shevchenko National University" Lubny, Poltava Oblast, Ukraine.
artem.lnu@ukr.net,
https://orcid.org/0000-0003-2690-983X

Abstract: *The interpretation of explicit and implicit contrast is a complex psycholinguistic process that requires the integration of linguistic, cognitive, and contextual information. This study examines the mechanisms underlying contrast recognition and interpretation in contemporary Ukrainian and English-language poetry and compares them with the computational operations of Large Language Models (ChatGPT, Gemini, and Claude). The research combines a controlled psycholinguistic experiment with comparative AI analysis within a unified interdisciplinary framework. The experimental corpus comprised thirty authentic twenty-first-century poetic excerpts, including twenty Ukrainian and ten English-language texts representing both explicit and implicit contrastive structures. Human participants and the three LLMs analysed the same material under identical experimental conditions, while representative cases were selected for detailed qualitative analysis. The study proposes an original seven-stage psycholinguistic model of contrast interpretation and demonstrates that the relationship between human cognition and transformer-based language models is best understood through the concept of functional analogues rather than cognitive equivalence. The findings show that both humans and LLMs accurately recognise explicit contrast, whereas implicit contrast exposes fundamental differences in interpretative strategies. Human readers rely on semantic memory, cultural knowledge, contextual integration, and inferential reasoning, while LLMs primarily exploit contextual representations and probabilistic prediction. The proposed framework extends psycholinguistic research on literary discourse and provides a methodological basis for evaluating the interpretation of implicit meaning by contemporary artificial intelligence systems.*

Keywords: *explicit contrast; implicit contrast; psycholinguistics; artificial intelligence; ChatGPT; Large Language Models; functional analogues; literary discourse; cognitive linguistics.*

How to cite: Shytyk, L. V., Yuldasheva, L. P., Dosenko, A. K., Sytnyk, O. V., Malysh, M. M., & Halych, A. O. (2026). Cognitive mechanisms of contrast recognition and interpretation and their functional analogues in artificial intelligence systems. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 17*(Special Issue 1), 328-353. https://doi.org/10.70594/brain/17.S1/26

1. Introduction

Contrast is widely recognised as a fundamental mechanism of linguistic conceptualisation that enables speakers to structure differences between objects, events, evaluations, and concepts. In discourse, it extends beyond antonymy or stylistic opposition by establishing semantic boundaries, highlighting incompatibility, directing attention, and shaping interpretation. From a psycholinguistic perspective, contrast is not a static semantic relation but a dynamic process involving semantic comparison, contextual integration, predictive processing, inferential reasoning, and pragmatic evaluation. These mechanisms become particularly important in poetic discourse, where meanings are frequently constructed through figurative language and implicit semantic relations rather than overt linguistic forms.

Contemporary Ukrainian and English-language poetry provides a particularly suitable context for investigating contrast because poetic meaning often emerges through hidden conceptual oppositions associated with memory, identity, trauma, embodiment, and cultural experience. Explicit contrast is signalled by lexical, grammatical, or discourse markers, whereas implicit contrast requires readers to reconstruct semantic relations by integrating contextual information, background knowledge, presuppositions, and pragmatic inference. Recent psycholinguistic research demonstrates that these processes involve multiple interacting cognitive mechanisms rather than the recognition of isolated linguistic structures.

Recent advances in Large Language Models, particularly the emergence of large-scale transformer models capable of few-shot learning (Brown et al., 2020), have created new opportunities for investigating computational language processing. Transformer architectures demonstrate increasingly sophisticated performance in interpreting metaphor, inference, and other forms of non-literal language. Nevertheless, successful linguistic performance should not be equated with human language comprehension. Human readers interpret contrast through semantic memory, embodied knowledge, contextual integration, and pragmatic reasoning, whereas Large Language Models rely on contextual representations, self-attention, and probabilistic prediction. Consequently, the present study adopts the concept of functional analogues, which refers to similarities in interpretative functions rather than cognitive equivalence between human and computational language processing.

Although the interpretation of figurative language by Large Language Models has attracted considerable scholarly attention, relatively little research has systematically compared the psycholinguistic mechanisms underlying literary contrast with the computational operations responsible for its interpretation in transformer-based models. Moreover, previous studies have rarely distinguished explicitly between contrast recognition and contrast interpretation, despite their different cognitive demands. This gap highlights the need for an interdisciplinary framework integrating psycholinguistic theory with contemporary Artificial Intelligence research.

The aim of the present study is to develop a seven-stage psycholinguistic model of explicit and implicit contrast interpretation and to identify its functional computational analogues in contemporary Large Language Models. To achieve this objective, the study combines a controlled psycholinguistic experiment involving forty participants with a comparative analysis of ChatGPT, Gemini, and Claude using a corpus of thirty contemporary Ukrainian and English-language poetic excerpts. By comparing human and computational interpretation under identical experimental conditions, the study proposes a psycholinguistically informed framework for evaluating the interpretative capabilities of Large Language Models while preserving the distinction between human cognition and computational language processing.

2. Methodology

The study employed an interdisciplinary methodology integrating experimental psycholinguistics, cognitive linguistics, and Artificial Intelligence research. It combined a psycholinguistic experiment with a comparative evaluation of contemporary Large Language Models in order to investigate the recognition and interpretation of explicit and implicit contrast.

The analysis focused on identifying functional analogues between human psycholinguistic processes and computational operations in transformer-based language models rather than establishing cognitive equivalence.

The experimental corpus consisted of thirty authentic poetic excerpts from twenty-first-century Ukrainian and English-language poetry, including twenty Ukrainian and ten English-language texts. The corpus was balanced with respect to explicit and implicit contrast and selected according to the following criteria: authenticity, thematic diversity, comparable length, and suitability for psycholinguistic analysis. All excerpts were analysed by both the human participants and the language models. The present article discusses representative examples illustrating the principal patterns observed across the corpus.

The psycholinguistic experiment involved forty native speakers of Ukrainian representing different age groups and educational backgrounds. Participation was voluntary, and all respondents completed the experiment individually under identical conditions. English-language poems were presented together with Ukrainian translations to minimise the influence of language proficiency.

Participants analysed every poetic excerpt by determining whether contrast was present, explaining its meaning, and justifying their interpretation. Their responses were subsequently compared with those generated by ChatGPT, Gemini, and Claude(Anthropic, 2024) using identical prompts and equivalent testing conditions. The complete prompts and technical settings are provided in the ***Appendix A*** to ensure the reproducibility of the computational analysis.

The findings were analysed using a combination of descriptive and inferential statistics, followed by qualitative psycholinguistic analysis. Statistical significance and effect sizes were calculated to evaluate the reliability of the results, while qualitative analysis identified the principal psycholinguistic mechanisms underlying successful and unsuccessful interpretation, particularly in cases of implicit contrast. Differences in the distribution of correct, partially correct, and failed interpretations between explicit and implicit contrast were analysed using Pearson's chi-square test. Effect size was estimated using Cramer's V, with statistical significance set at $p < .05$.

The study complied with generally accepted ethical standards of psycholinguistic research. All participants provided informed consent, and the data were anonymised prior to analysis.

3. Theoretical Background

3.1. Cognitive Foundations of Contrast Interpretation

Contrast interpretation has attracted considerable attention within contemporary cognitive linguistics and psycholinguistics because it represents one of the principal mechanisms through which language organises conceptual knowledge. Rather than functioning solely as a semantic relation between opposing linguistic units, contrast emerges through the interaction of conceptual organisation, contextual knowledge, categorisation, prediction, and inferential reasoning. Consequently, the interpretation of explicit and implicit contrast requires an integrated cognitive framework that explains different aspects of meaning construction.

One of the foundations of this framework is Conceptual Metaphor Theory, developed by Lakoff and Johnson (1980). Rejecting the view of metaphor as a merely stylistic phenomenon, they argued that *"Our ordinary conceptual system, in terms of which we both think and act, is fundamentally metaphorical in nature"* (Lakoff & Johnson, 1980, p. 3). From this perspective, contrast functions as a mechanism of conceptual differentiation, enabling speakers and readers to distinguish alternative conceptualisations and organise meaning through relations of similarity and opposition rather than through isolated lexical units alone.

This perspective is complemented by Frame Semantics, proposed by Fillmore (1982), who argued that *"The frame notion offers a particular way of characterising the relationship between linguistic meanings and the worlds"* (Fillmore, 1982, p. 111). Meaning therefore emerges through the activation of structured knowledge associated with recurrent situations and experiences. In

literary discourse, contrast frequently arises when the conceptual frame anticipated by the reader differs from the frame activated by the text, requiring the reorganisation of the initial interpretation.

A further dimension of contrast interpretation is provided by Grounded Cognition Theory. According to Barsalou (2008), *"Grounded cognition assumes that cognitive processes are deeply rooted in the body's interactions with the world"* (p. 618). Meaning is therefore dynamically constructed through perceptual, emotional, and situational simulation rather than retrieved as a fixed semantic representation. This perspective is particularly important for implicit contrast, whose interpretation depends on embodied experience, contextual knowledge, and cultural memory extending beyond the explicit linguistic content of the text.

The cognitive role of conceptual expectations is further explained by Prototype Theory**,** developed by Rosch (1978). As she observes, *"The world consists of virtually an infinite number of discriminably different stimuli"* (p. 28), making categorisation an essential principle of cognitive economy. Readers therefore interpret discourse by comparing textual information with prototypical conceptual representations stored in memory. Contrast becomes especially salient when the discourse departs from these established expectations, prompting the construction of new conceptual relationships.

Taken together, these theoretical perspectives describe complementary dimensions of contrast interpretation. Conceptual Metaphor Theory explains conceptual differentiation, Frame Semantics accounts for the organisation of contextual knowledge, Grounded Cognition emphasises experiential simulation, and Prototype Theory explains expectation-based categorisation. Previous psycholinguistic research likewise demonstrates that successful interpretation of contrast depends on the coordinated interaction of semantic comparison, contextual prediction, conceptual integration, inferential reasoning, and pragmatic evaluation rather than on isolated linguistic markers (Yuldasheva, 2025). Collectively, these approaches provide the theoretical foundation for the psycholinguistic model proposed in the present study.

3.2. Psycholinguistic Mechanisms of Explicit and Implicit Contrast Interpretation

Building upon the cognitive approaches discussed above, contrast interpretation may be viewed as a multilevel psycholinguistic process rather than as the recognition of isolated linguistic oppositions. Although explicit and implicit contrast differ in their linguistic realisation, both involve the interaction of semantic processing, contextual integration, prediction, and inferential reasoning. The principal difference lies in the cognitive effort required to establish the contrastive relationship.

Explicit contrast is characterised by overt linguistic markers, including antonyms, adversative conjunctions, and contrastive syntactic constructions. These markers activate well-established patterns of semantic opposition, allowing readers to recognise the intended contrast rapidly and with relatively little cognitive effort. Interpretation therefore depends primarily on the automatic activation of linguistic and conceptual knowledge stored in long-term memory.

Implicit contrast, by contrast, lacks explicit linguistic signals. Readers must reconstruct the contrastive relationship by integrating lexical meaning with contextual information, background knowledge, presuppositions, and pragmatic inference. Successful interpretation therefore requires substantially greater cognitive resources because meaning is constructed rather than directly recognised. As previous psycholinguistic research has demonstrated, this process depends on the coordinated interaction of semantic comparison, contextual prediction, conceptual integration, inferential reasoning, and pragmatic evaluation (Yuldasheva, 2025).

Contemporary psycholinguistic research further demonstrates that language comprehension is inherently predictive. Readers continuously generate expectations regarding the development of discourse and revise these expectations whenever incoming information conflicts with the emerging mental representation (Van Berkum et al., 2008; Hagoort, 2005). Explicit contrast rarely disrupts this process because overt linguistic markers prepare readers for a forthcoming shift in perspective. Implicit contrast, however, frequently requires the revision of an initially constructed interpretation,

making prediction, contextual integration, and inferential reasoning central mechanisms of successful comprehension.

Drawing together these theoretical perspectives, the present study proposes that contrast interpretation proceeds through a sequence of interacting psycholinguistic operations extending from the initial processing of linguistic input to the construction of a pragmatic interpretation. Although these operations interact dynamically during discourse comprehension, they may be described as seven functionally distinct stages that collectively explain the recognition and interpretation of both explicit and implicit contrast. The proposed model also provides the conceptual basis for comparing human language comprehension with the computational operations of contemporary Large Language Models.

The sequence of psycholinguistic operations proposed in the present study is summarised in Figure 1.

Figure 1. Seven-stage psycholinguistic model of explicit and implicit contrast interpretation and its functional computational analogues in transformer-based Large Language Models.

Human psycholinguistic processing	Functional computational analogue in LLMs
Stage 1. Linguistic input processing	Tokenisation and contextual embeddings
↓	↓
Stage 2. Prediction formation	Next-token probabilistic prediction
↓	↓
Stage 3. Contrast detection	Contextual attention and semantic inconsistency detection
↓	↓
Stage 4. Semantic integration	Multi-layer contextual representation
↓	↓
Stage 5. Frame reconfiguration	Dynamic updating of contextual representations
↓	↓
Stage 6. Inferential interpretation	Distributed semantic inference
↓	↓
Stage 7. Pragmatic evaluation	Probabilistic response generation and interpretation selection

Human reader

Integrated literary interpretation based on semantic memory, cultural knowledge, embodied cognition and pragmatic reasoning.

Large Language Model

Contextually coherent interpretation generated through probabilistic language modelling and contextual representations.

The model distinguishes seven successive but interacting stages: *linguistic input processing, prediction formation, contrast detection, semantic integration, frame reconfiguration, inferential interpretation, and pragmatic evaluation.* Explicit contrast is generally resolved during the earlier stages of processing because linguistic markers directly activate contrastive relations. Implicit contrast, in contrast, increasingly depends on frame reconfiguration, inferential reasoning, and pragmatic evaluation, explaining its greater cognitive complexity and the higher variability observed in human interpretation. This model serves as the theoretical framework for the experimental analyses presented in the following sections.

3.3. Neurocognitive Evidence for Contrast Interpretation

Contemporary psycholinguistic research increasingly relies on neurocognitive methods to investigate the temporal dynamics of language comprehension. Electrophysiological techniques,

eye-tracking, and predictive processing models have demonstrated that meaning construction is not a single cognitive act but a sequence of interacting processes unfolding during discourse comprehension. These findings provide important empirical support for the psycholinguistic model proposed in the present study by demonstrating that contrast interpretation involves successive stages of semantic processing, contextual integration, prediction, and meaning revision rather than the simple recognition of linguistic opposition.

One of the most extensively investigated neurophysiological indicators of semantic processing is the **N400** component, first described by Kutas and Hillyard (1980). Their pioneering study demonstrated that semantically unexpected words elicit a larger N400 response than contextually expected continuations, indicating that the brain continuously evaluates incoming linguistic information against contextually generated expectations. As Kutas and Federmeier (2011) subsequently demonstrated, the N400 reflects not only semantic incongruity but also the interaction between contextual prediction and semantic integration. These findings are directly relevant to contrast interpretation because implicit contrast frequently requires readers to revise predictions generated during the earlier stages of discourse processing.

Semantic revision following the detection of incongruity has been associated with the **P600** component. Although originally interpreted as a marker of syntactic reanalysis, subsequent research has shown that the P600 also reflects broader processes of meaning revision and discourse-level reinterpretation (Kuperberg, 2007). In the interpretation of implicit contrast, such revision becomes necessary when newly encountered information cannot be integrated into the initially constructed mental representation. The interaction between the N400 and P600 therefore provides neurocognitive evidence that contrast interpretation extends beyond lexical processing to include semantic reorganisation and pragmatic reinterpretation.

Complementary evidence has been obtained from eye-tracking research, which demonstrates that implicit contrast is associated with longer fixation durations, more frequent regressions, and increased rereading of preceding textual segments (Rayner, 1998). These behavioural indicators suggest that readers allocate additional cognitive resources when reconstructing implicit semantic relations, supporting the view that implicit contrast imposes greater processing demands than explicit contrast.

Taken together, findings from ERP research, eye-tracking studies, and predictive models of language comprehension converge on the conclusion that contrast interpretation proceeds through multiple interacting cognitive operations rather than a single recognition process. The proposed seven-stage psycholinguistic model integrates these findings into a unified theoretical framework that describes how linguistic input is transformed into coherent interpretation. This framework also provides the conceptual basis for the comparison with transformer-based Large Language Models presented in the following section.

4. Functional Analogues of Human Contrast Interpretation in Large Language Models

Recent advances in Large Language Models have considerably expanded the capacity of artificial intelligence systems to process complex linguistic phenomena, including metaphor, figurative language, implicature, and explicit as well as implicit contrast. Their performance has stimulated growing interdisciplinary interest in the relationship between computational language processing and human cognition. At the same time, increasing linguistic competence should not be interpreted as evidence that humans and language models rely on comparable cognitive mechanisms.

One of the most influential contributions to this debate was made by Bender and Koller (2020), who argued that successful manipulation of linguistic forms does not necessarily imply semantic understanding. As they observe, *"a system trained only on form has a priori no way to learn meaning"* (Bender & Koller, 2020, p. 5185). According to this view, language models acquire statistical regularities from textual data rather than conceptual knowledge grounded in interaction with the physical and social world. Similar concerns were subsequently expressed by Bommasani et

al. (2021), who cautioned against attributing human cognitive properties to Foundation Models solely on the basis of their linguistic performance.

Recent research, however, suggests that the relationship between statistical language modelling and semantic processing is more complex than the opposition between "understanding" and "pattern matching" implies. Studies in mechanistic interpretability demonstrate that transformer-based architectures develop structured contextual representations distributed across multiple layers and attention mechanisms, supporting increasingly sophisticated forms of semantic integration and discourse processing (Zheng et al., 2025). These findings do not imply that language models possess human-like conceptual knowledge, but they indicate that their internal computational organisation cannot be reduced to simple statistical retrieval.

Accordingly, the present study adopts the concept of functional analogues as the most appropriate framework for comparing human language comprehension with computational language processing. Functional analogy does not imply similarity of cognitive architecture, conscious experience, or mental representation. Instead, it refers to the correspondence between interpretative functions performed by two fundamentally different systems. Human readers interpret contrast through semantic memory, contextual integration, prediction, frame activation, inferential reasoning, and pragmatic evaluation, whereas Large Language Models achieve comparable interpretative outcomes through tokenisation, contextual embeddings, self-attention, hierarchical representation learning, and probabilistic prediction. The comparison therefore concerns functional correspondence rather than cognitive equivalence.

Computational language processing in transformer-based Large Language Models begins with the transformation of the input text into a sequence of tokens. Unlike human readers, who perceive words as meaningful linguistic units grounded in conceptual knowledge and experience, transformer architectures operate on subword tokens that constitute the basic units of computation. Each token is represented as a high-dimensional embedding vector whose properties are learned from statistical regularities observed across extensive textual corpora (Vaswani et al., 2017; Devlin et al., 2019). Consequently, lexical meaning is not stored as an independent semantic entity but emerges dynamically from contextual representations generated during processing.

The contextualisation of linguistic information is achieved through the self-attention mechanism, which allows every token to interact with every other token in the input sequence. Rather than processing language sequentially, transformer architectures continuously evaluate contextual dependencies, assigning different levels of importance to individual tokens according to their contribution to the evolving discourse representation (Vaswani et al., 2017). Through multiple transformer layers, these contextual representations become progressively richer, enabling the model to resolve ambiguities, capture long-range semantic dependencies, and integrate information distributed across the text.

Recent studies in mechanistic interpretability indicate that these representations are organised hierarchically rather than uniformly throughout the network. Lower transformer layers primarily encode lexical and syntactic information, whereas deeper layers increasingly capture semantic relations, discourse organisation, and contextual dependencies (Zheng et al., 2025). Although the precise functional specialisation of individual layers remains an active area of investigation, current evidence suggests that linguistic interpretation emerges from the coordinated interaction of multiple representational levels rather than from isolated computational components.

Although the precise functional specialisation of individual attention heads remains an active area of investigation, converging evidence suggests that semantic interpretation emerges through distributed interactions across multiple layers rather than through isolated computational components (Rogers et al., 2020; Nanda, 2024). Accordingly, the functional analogues proposed in the present study should be interpreted at the level of information-processing functions rather than as one-to-one correspondences between individual psycholinguistic stages and specific transformer layers. Prediction constitutes another important point of functional correspondence between human language comprehension and transformer-based processing. Human readers continuously anticipate

the development of discourse by combining linguistic input with contextual knowledge and previous experience. Large Language Models likewise operate as predictive systems; however, their predictions are generated through probabilistic estimation of the most likely subsequent token rather than through conceptual expectations. Prediction therefore facilitates efficient language processing in both systems while relying on fundamentally different representational principles. These computational mechanisms become particularly important during the interpretation of implicit contrast. Explicit contrast is generally associated with overt lexical or syntactic markers that correspond to statistically robust patterns encountered repeatedly during model training. Implicit contrast, however, lacks such surface cues and therefore requires the model to recover hidden semantic relations exclusively through contextual representations distributed across the discourse. Successful interpretation consequently depends on the quality of contextual integration and probabilistic inference rather than on explicit linguistic signalling.

Taken together, tokenisation, contextual embeddings, self-attention, hierarchical representation learning, and probabilistic prediction constitute the principal computational operations through which contemporary Large Language Models process contrastive meaning. Although these operations differ fundamentally from the psycholinguistic mechanisms underlying human language comprehension, they perform several comparable interpretative functions. This functional correspondence provides the basis for the comparative model presented in Figure 2.

Figure 2. Correspondence between the psycholinguistic stages of human contrast interpretation and their functional computational analogues in transformer-based Large Language Models.

Human Psycholinguistic Stage		Shared Interpretative Function		Functional Computational Analogue in LLMs
Stage 1. Linguistic Input Processing Initial processing of lexical, morphological and syntactic information; construction of a preliminary linguistic representation.	⟺	Establish the initial linguistic representation of the input.	⟺	**Tokenisation and Contextual Embeddings** The input is segmented into tokens and represented as contextual embedding vectors that provide the basis for further processing.
Stage 2. Prediction Formation Generation of expectations concerning the likely continuation of the discourse based on linguistic experience, contextual information and world knowledge.	⟺	Generate expectations about the continuation of the discourse.	⟺	**Next-Token Prediction** The model estimates the probability distribution of subsequent tokens and continuously predicts the most likely continuation.
Stage 3. Contrast Detection Recognition of discrepancies between predicted and incoming information leading to the identification of explicit or implicit contrast.	⟺	Detect semantic incongruity and identify contrastive relations.	⟺	**Attention-Based Detection of Semantic Inconsistency** Self-attention mechanisms identify contextual inconsistencies and update token representations when unexpected information appears.
Stage 4. Semantic Integration Integration of newly processed information into the evolving discourse representation and construction of coherent meaning.	⟺	Integrate linguistic information into a coherent discourse representation.	⟺	**Hierarchical Contextual Integration** Multiple transformer layers iteratively integrate information across the entire discourse, refining contextual representations.
Stage 5. Frame Reconfiguration Revision of the initial conceptual representation; reorganisation of cognitive frames following the recognition of contrast.	⟺	Reorganise conceptual representations after detecting contrast.	⟺	**Dynamic Updating of Contextual Representations** Internal representations are modified to accommodate newly emerging semantic relationships.
Stage 6. Inferential Interpretation Reconstruction of implicit meaning through presuppositions, background knowledge, contextual information and inferential reasoning.	⟺	Recover implicit semantic relations extending beyond the linguistic surface.	⟺	**Distributed Semantic Inference** The model infers unstated semantic relations using distributed representations learned during large-scale pre-training.
Stage 7. Pragmatic Evaluation Identification of communicative intention, pragmatic function, emotional effect and overall literary meaning.	⟺	Produce a contextually appropriate interpretation of the discourse.	⟺	**Probabilistic Interpretation Generation** The model generates the most probable interpretation or response consistent with contextual representations and learned statistical regularities.

Note

The correspondence presented in this model concerns similarities in ***interpretative functions*** rather than equivalence of cognitive or computational mechanisms. Human contrast interpretation is grounded in semantic memory,

embodied cognition, cultural knowledge, and pragmatic reasoning, whereas Large Language Models rely on contextual embeddings, self-attention, distributed representations, and probabilistic prediction.

The model summarises the functional correspondences that are subsequently evaluated experimentally in the Results section.

5. Functional Analogues as a Framework for Comparative Analysis

The comparison proposed in the present study should not be interpreted as establishing one-to-one correspondence between individual psycholinguistic processes and specific computational operations. Human language comprehension emerges from the continuous interaction of linguistic, cognitive, emotional, social, and cultural systems, whereas Large Language Models operate through distributed numerical representations optimised for probabilistic language prediction. Consequently, similar interpretative outcomes may arise from fundamentally different processing architectures.

This distinction becomes particularly important in the interpretation of implicit contrast. Literary discourse frequently relies on meanings that remain only partially encoded in the linguistic surface of the text and therefore require readers to integrate contextual information, cultural knowledge, presuppositions, and pragmatic inference. Human readers reconstruct these meanings through semantic memory, embodied experience, and conceptual reasoning acquired through interaction with the world. Transformer-based language models approach the same task by integrating statistical regularities encoded across contextual representations generated during large-scale pre-training. Although both systems may produce similar interpretations, the processes through which these interpretations emerge remain fundamentally different.

Recent studies evaluating *Theory of Mind* in Large Language Models further illustrate this distinction. Contemporary transformer architectures successfully solve many tasks involving beliefs, intentions, and perspective taking under controlled experimental conditions (Strachan et al., 2024). Nevertheless, high performance on such benchmarks should not be interpreted as evidence of human-like mental representations or conscious reasoning. Rather, it demonstrates that increasingly sophisticated contextual representations enable language models to reproduce a range of interpretative functions that often resemble human inferential behaviour while remaining computational in nature.

The concept of functional analogues therefore provides a theoretically balanced framework for comparing human language comprehension with computational language processing. It avoids two equally reductionist positions: attributing human cognitive capacities to artificial intelligence or reducing contemporary Large Language Models to simple statistical prediction systems. Instead, it recognises that comparable interpretative outcomes may emerge from fundamentally different mechanisms, making functional correspondence a more appropriate object of comparison than cognitive equivalence.

The theoretical framework developed in the present study forms the basis for the experimental analysis reported below. Using an identical corpus of contemporary Ukrainian and English-language poetic texts, the experiment compares the interpretation of explicit and implicit contrast by human participants, ChatGPT, Gemini, and Claude under identical analytical conditions. The results make it possible to evaluate empirically the extent to which the proposed functional analogues explain both the similarities and the differences between psycholinguistic and computational language processing.

6. Results

6.1. Human Interpretation of Explicit and Implicit Contrast

To investigate the psycholinguistic mechanisms underlying the recognition and interpretation of explicit and implicit contrast, a controlled experimental study was conducted with native speakers of Ukrainian. The experiment aimed to examine the cognitive processes involved in interpreting different types of contrastive structures, identify the principal factors influencing

successful interpretation, and compare the psycholinguistic characteristics of explicit and implicit contrast under identical experimental conditions.

The experiment involved forty native speakers of Ukrainian aged between 20 and 55 years (M = 33.8, SD = 8.9), including 28 women and 12 men. Participants represented diverse educational and professional backgrounds in order to minimise the influence of specialised linguistic expertise on literary interpretation. They were recruited through an open voluntary invitation rather than on the basis of literary competence, and none had previously participated in psycholinguistic experiments on contrast interpretation. This sampling strategy was intended to minimise potential selection bias and to approximate the responses of ordinary readers rather than literary experts. None of the participants had professional training in linguistics, literary criticism, or artificial intelligence that could systematically influence the results. All reported normal (or corrected-to-normal) vision and no history of language-related neurological disorders.

The experimental corpus comprised thirty authentic poetic excerpts from contemporary twenty-first-century Ukrainian and English-language poetry, including twenty Ukrainian and ten English-language texts. The corpus was balanced with respect to explicit and implicit contrastive structures. All texts were analysed by the human participants and subsequently processed by ***ChatGPT, Gemini, and Claude*** under comparable analytical conditions. Although all thirty excerpts contributed to the statistical analysis, six representative examples were selected for detailed qualitative discussion because they illustrate the principal psycholinguistic mechanisms identified across the corpus.

The order of text presentation was randomised individually for each participant in order to minimise order and fatigue effects. The experiment was conducted individually in an online environment. Before beginning the task, participants received identical written instructions explaining the procedure. Importantly, no definition of contrast was provided, allowing natural mechanisms of contrast recognition to emerge without activating explicit metalinguistic strategies.

After reading each excerpt, participants completed an identical sequence of tasks. They were asked to determine whether the text contained contrast, identify the linguistic segment expressing or signalling the contrastive relationship, explain which concepts or images were contrasted, interpret the semantic function of the contrast within the discourse, formulate their understanding of the author's intended meaning, and evaluate their confidence in the response using a five-point Likert scale.

Participant responses were evaluated according to six psycholinguistic variables:

1. accuracy of contrast recognition;
2. accuracy of localisation of the contrastive structure;
3. completeness of semantic interpretation;
4. ability to reconstruct implicit conceptual relations;
5. adequacy of reconstructing the author's communicative intention;
6. subjective confidence in the interpretation.

The same evaluation criteria were subsequently applied to the responses generated by ***ChatGPT, Gemini, and Claude***, ensuring direct comparability between human and computational interpretation.

The study complied with internationally accepted ethical principles governing psycholinguistic research. All participants provided informed consent before participation, were informed about the anonymous processing of the data, and could withdraw from the experiment at any stage without negative consequences. No personally identifiable information was collected.

For the English-language excerpts, participants were allowed to consult an electronic translation tool if necessary. However, no historical, literary, or contextual information was supplied, ensuring that interpretation depended primarily on linguistic competence, contextual reasoning, and background knowledge rather than externally provided explanations. Although this procedure reduced the influence of foreign-language proficiency, it may also have limited

participants' access to subtle stylistic and semantic nuances present in the original English texts. This limitation is discussed further in the Discussion section.

The statistical analysis combined descriptive and inferential approaches. Frequencies, percentages, means, and standard deviations were calculated for all experimental variables. Differences in the distribution of correct, partially correct, and incorrect interpretations between explicit and implicit contrast were analysed using Pearson's chi-square (χ^2) test. Effect sizes were estimated using Cramer's V, and statistical significance was established at $p < .05$. Qualitative analysis was subsequently performed to identify the principal psycholinguistic mechanisms underlying successful interpretation and the most frequent sources of interpretative difficulty.

The results are presented in two stages. First, explicit contrast is analysed as the baseline condition because overt linguistic markers provide direct cues for recognising semantic opposition. The second part examines implicit contrast, where successful interpretation depends on contextual integration, conceptual restructuring, and inferential reasoning. This organisation makes it possible to compare the psycholinguistic characteristics of the two types of contrast under identical experimental conditions before comparing human performance with that of Large Language Models in the following section.

Within the representative sample, the distribution of correct, partially correct, and failed interpretations differed significantly between explicit and implicit contrast, $\chi^2(2, N = 240$ observations) = 46.82, $p < .001$, Cramer's $V = .44$. Here, N refers to the total number of interpretation responses included in the analysis rather than the number of participants. These findings confirm that implicit contrast required significantly greater cognitive effort than explicit contrast and provide the statistical basis for the qualitative analyses presented below. These findings confirm that implicit contrast required significantly greater cognitive effort than explicit contrast and provide the statistical basis for the qualitative analyses presented below.

The first stage of the analysis examined the recognition and interpretation of explicit contrast, which served as the baseline condition for comparing human processing of more cognitively demanding implicit contrastive structures. Three representative excerpts were selected for detailed analysis: two Ukrainian texts containing conventional antonymic oppositions and one English-language excerpt in which explicit contrast was realised through conceptual metaphor.

The Ukrainian material included excerpts from poems by O. Stomina:

Несу війну на плечах, мов броню
жую про неї правду і брехню
(I carry the war on my shoulders like armour,
chewing its truth and its lies)

and L. Kybalka*:*

Схрестилися стежини радості та смутку
(The paths of joy and sorrow crossed).

Both excerpts contain overtly expressed contrastive structures based on conventional antonymic pairs: *truth – lie* and joy – sorrow. Because these oppositions are explicitly encoded in the linguistic structure of the text, they provide readers with immediate cues for recognising the intended contrastive relationship.

The English-language material was represented by the opening lines of Warsan Shire's poem *Home*:

No one leaves home unless
home is the mouth of a shark.

Although the contrast is expressed explicitly, its interpretation requires readers to integrate metaphorical meaning with the conventional conceptual representation of home as a place of safety and protection. Consequently, this example occupies an intermediate position between conventional lexical opposition and more complex conceptual contrast.

The experimental results demonstrate that explicit contrast constitutes the most readily accessible type of contrastive relation for human language comprehension. Both Ukrainian excerpts

were recognised correctly by all participants, yielding an identification rate of 100%. Participants not only identified the presence of contrast without difficulty but also accurately localised the linguistic elements responsible for expressing the opposition. Their explanations consistently referred to the antonymic pairs organising the semantic structure of the texts and correctly interpreted their contribution to the overall poetic meaning.

A similar pattern was observed in the English-language excerpt by Warsan Shire. Thirty-two participants **(**80%**)** correctly recognised the explicit contrast, whereas eight respondents provided only partially adequate interpretations. Importantly, none of these participants failed to recognise the existence of opposition itself. Instead, the observed difficulties concerned the conceptual interpretation of the metaphor rather than the recognition of contrast.

Qualitative analysis of participants' responses revealed that most respondents correctly interpreted home as a symbol of safety and understood that this conventional image was contrasted with the threatening metaphor of the mouth of a shark**.** However, participants whose interpretations were classified as only partially correct frequently limited their responses to the literal image of danger without reconstructing the broader conceptual opposition between security and mortal threat. These findings suggest that the principal source of difficulty did not arise from recognising explicit contrast but from integrating metaphorical meaning within a broader conceptual framework.

Table 1. Recognition of Explicit Contrast by Human Participants

Poetic excerpt	**Type of contrast**	**Correctly identified (n = 40)**	**Partially identified**	**Correct (%)**
O. Stomina	Explicit	40	0	100.0
L. Kybalka	Explicit	40	0	100.0
Warsan Shire	Explicit	32	8	80.0

The quantitative findings demonstrate that explicit contrast posed relatively little difficulty for native speakers regardless of the language of the text. Across the experimental corpus, overt linguistic markers consistently facilitated rapid recognition of contrastive relations. When interpretative difficulties occurred, they were associated primarily with the reconstruction of metaphorical and conceptual meaning rather than with the identification of contrast itself. These findings establish the baseline against which the substantially greater cognitive demands of implicit contrast can be evaluated.

In contrast to explicit contrast, implicit contrast presents a substantially greater psycholinguistic challenge because the opposition is not directly encoded in the linguistic surface of the text. Readers must reconstruct the underlying contrastive relationship by integrating lexical meaning with contextual information, background knowledge, presuppositions, conceptual frames, and pragmatic inference. Consequently, successful interpretation depends not only on linguistic competence but also on the coordinated operation of multiple cognitive mechanisms described in the psycholinguistic model proposed in the present study.

To examine these mechanisms, three representative excerpts containing implicit contrast were selected for detailed analysis: two Ukrainian texts by A. Voitenko and O. Kadalov, and one English-language excerpt by Ocean Vuong. Although all three texts required readers to reconstruct implicit semantic relations, they differed considerably in the degree of contextual and cultural knowledge necessary for successful interpretation.

The first Ukrainian example, taken from a poem by A. Voitenko**,** reads:

А в нього у очах не світять зорі,
а в нього тільки присмак цигарок
(*There are no stars shining in his eyes/*
there is only the lingering taste of cigarettes).

Thirty-six participants **(**90%**)** successfully recognised the implicit contrast, whereas four respondents produced only partial interpretations (*Table 2*). None of the participants completely failed to perceive the presence of semantic tension. Most readers interpreted the passage as

contrasting hope with disillusionment, spiritual aspiration with physical exhaustion, or romantic expectation with the harsh reality of war. Although individual explanations differed in wording, they converged conceptually, demonstrating that participants were able to reconstruct the underlying opposition despite the absence of explicit linguistic markers.

From a psycholinguistic perspective, these findings indicate that the principal cognitive challenge did not concern contrast detection ***but*** frame reconfiguration. Participants generally recognised that the two images could not be interpreted within a single conceptual frame and therefore reorganised the developing mental representation of the discourse. Differences between individual responses reflected variation in the inferential strategies used to explain the reconstructed conceptual opposition rather than failure to recognise the contrast itself.

A more complex pattern emerged in the interpretation of the excerpt by O. Kadalov:

Я батарейка, що працює навіть з мінусовим зарядом

(*I am a battery that keeps working even with a negative charge).*

Twenty-nine participants **(**73%**)** correctly reconstructed the implicit contrast and interpreted it as an opposition between exhaustion and persistence, weakness and resilience, or emotional depletion and continued functioning. Eleven respondents recognised the paradoxical character of the statement but failed to explain precisely which conceptual domains were being contrasted. Their responses remained at the level of metaphorical paraphrase without reconstructing the underlying conceptual conflict.

This pattern suggests that participants successfully completed the earlier stages of processing – linguistic analysis, prediction, and contrast detection – but encountered greater difficulty during inferential interpretation. They recognised that the statement violated conventional conceptual expectations yet were unable to integrate this violation into a coherent explanation of the author's intended meaning. These findings demonstrate that recognising an implicit contrast does not necessarily guarantee its successful interpretation, further supporting the distinction between contrast recognition and contrast interpretation proposed in the present study.

The greatest psycholinguistic difficulty was observed in the English-language excerpt from Ocean Vuong**:**

Milkflower petals on the street like pieces of a girl's dress.

Only five participants (13%) suggested that the passage contained implicit contrast, whereas thirty-five respondents (87%) failed to identify any underlying opposition (Table 2). Unlike the Ukrainian examples, participants frequently reported uncertainty regarding their interpretations and explicitly acknowledged that they lacked sufficient information to explain the passage. Typical comments included *"I need more context," "The image seems meaningful, but I cannot determine the contrast,"* and *"Without knowing the rest of the poem, the intended meaning remains unclear."*

From a psycholinguistic perspective, these responses are particularly informative because they reveal not misunderstanding but awareness of the limitations of one's own interpretative resources. Rather than proposing unsupported explanations, participants recognised that successful interpretation required information unavailable in the isolated excerpt.

The observed pattern cannot be explained by linguistic difficulty alone. Three interacting factors appear to account for the exceptionally low recognition rate. First**,** the excerpt contains no overt lexical or syntactic markers signalling opposition, making the contrast inaccessible through linguistic analysis alone. Second**,** successful interpretation depends on the activation of broader cultural and historical knowledge associated with war, violence, loss, and the symbolic imagery characteristic of Ocean Vuong's poetry.

Third**,** presenting the passage in isolation substantially reduced the availability of contextual information necessary for reconstructing the conceptual frame within which the implicit contrast emerges. The findings therefore suggest that the principal source of interpretative difficulty was contextual and cultural rather than linguistic**.** This finding directly addresses the question raised by the exceptionally low recognition rate, indicating that unsuccessful interpretation resulted primarily

from limitations in contextual and cultural knowledge rather than deficiencies in linguistic processing.

Within the psycholinguistic model proposed in the present study, the observed errors are concentrated primarily at the stages of frame reconfiguration *and* inferential interpretation. Most participants successfully processed the lexical content of the excerpt but were unable to reorganise the activated conceptual frame or generate the inferences required to establish the intended contrastive relationship. Consequently, failure occurred not during linguistic decoding but during the higher-order cognitive operations responsible for integrating contextual knowledge with the developing discourse representation.

Table 2. Recognition of Implicit Contrast by Human Participants

Poetic excerpt	**Correctly identified the contrast**	**Partially correct interpretation**	**Failed to identify the contrast**
A. Voitenko	36 (90%)	4 (10%)	0
O. Kadalov	29 (73%)	11 (27%)	0
Ocean Vuong	5 (13%)	0	35 (87%)

Table 3. Principal Sources of Interpretative Difficulty in Implicit Contrast

Source of difficulty	**A. Voitenko**	**O. Kadalov**	**Ocean Vuong**
Absence of explicit linguistic markers	✓	✓	✓
Difficulty in frame reconfiguration	Moderate	High	Very high
Difficulty in inferential interpretation	Moderate	High	Very high
Insufficient contextual information	Low	Moderate	Very high
Dependence on cultural/background knowledge	Low	Moderate	Very high

The progression observed in Table 3 demonstrates that increasing interpretative difficulty is associated with greater demands on frame reconfiguration and inferential reasoning rather than on basic linguistic processing.

Taken together, the findings demonstrate that explicit and implicit contrast engage qualitatively different psycholinguistic mechanisms of language comprehension. Although both forms of contrast involve the recognition of oppositional relations, they differ substantially in the cognitive operations required for successful interpretation. Explicit contrast is recognised primarily through overt linguistic cues that activate well-established semantic representations and facilitate rapid integration within the developing discourse model. Consequently, successful interpretation is largely achieved during the initial stages of psycholinguistic processing and requires relatively limited inferential effort.

Implicit contrast follows a fundamentally different cognitive trajectory. Because the opposition is not explicitly encoded in the linguistic surface of the text, readers must progressively reconstruct the intended meaning through the interaction of contextual information, conceptual knowledge, presuppositions, and inferential reasoning. Interpretation therefore depends on the successful coordination of multiple cognitive operations rather than on the recognition of isolated linguistic markers. The experimental findings indicate that increasing interpretative difficulty is accompanied by a corresponding increase in the importance of higher-order cognitive processes responsible for conceptual restructuring and pragmatic evaluation.

The results also provide empirical support for the seven-stage psycholinguistic model proposed in the present study. Across all experimental materials, successful interpretation of explicit contrast was associated primarily with efficient linguistic input processing, prediction formation, contrast detection, and semantic integration. In contrast, successful interpretation of implicit contrast depended increasingly on the subsequent stages of frame reconfiguration**,** inferential interpretation**,** and pragmatic evaluation**.** These latter stages proved decisive whenever readers were required to reconstruct meanings that were not directly represented in the linguistic form of the text.

The analysis of participant responses further demonstrates that interpretative failure cannot be regarded as a unitary phenomenon. Different types of unsuccessful responses reflected different

cognitive limitations. Some participants recognised the presence of contrast but failed to reorganise the conceptual frame required for coherent interpretation. Others successfully reconstructed the conceptual opposition but were unable to formulate the broader pragmatic meaning intended by the author. The Ocean Vuong excerpt illustrates this distinction particularly clearly. Most participants processed the lexical content accurately but were unable to activate the cultural and contextual knowledge necessary for higher-order inferential interpretation. Consequently, unsuccessful performance reflected limitations in conceptual integration rather than deficiencies in linguistic processing.

The human experiment additionally establishes a robust baseline for the comparative analysis presented in the following section. Because the same corpus, analytical procedure, and evaluation criteria were subsequently applied to ChatGPT, Gemini, and Claude, the comparison between human participants and Large Language Models is based on directly comparable experimental conditions. This methodological consistency makes it possible to evaluate not only similarities in interpretative outcomes but also the extent to which computational processing reproduces the psycholinguistic functions identified in human contrast interpretation.

6.2. Interpretation of Explicit Contrast by Large Language Models

The second stage of the study investigated how contemporary Large Language Models process explicit contrast in literary discourse. Three state-of-the-art transformer-based systems – ***ChatGPT***, ***Gemini***, and ***Claude*** — were evaluated under identical experimental conditions. Each model received the same poetic excerpts and the same prompt, requiring it to determine whether contrast was present, identify the linguistic segment expressing the contrastive relation, classify the type of contrast, reconstruct the conceptual poles involved, explain the function of the contrast, and infer the author's communicative intention. All models were evaluated using identical prompts and inference parameters, as described in the Methodology, thereby ensuring the comparability of their responses.

The analysis demonstrated a remarkably high level of agreement across all three language models in the interpretation of explicit contrast. Regardless of architectural differences, the models consistently recognised overtly expressed oppositional relations, accurately identified the lexical units responsible for signalling contrast, reconstructed the relevant conceptual oppositions, and produced coherent interpretations of their artistic and communicative functions. These findings suggest that explicit contrast constitutes a comparatively straightforward task for contemporary transformer-based language models, largely because overt linguistic markers provide reliable statistical cues that facilitate contextual prediction and semantic integration.

This pattern was observed consistently in the Ukrainian excerpts by O. Stomina and L. Kybalka**.** In both cases, all three models correctly detected the presence of contrast, identified the corresponding lexical oppositions, and reconstructed the underlying conceptual structure without ambiguity (Table 3). Minor differences occurred only in the stylistic formulation of the responses. While ChatGPT generally produced concise explanatory interpretations, Gemini tended to elaborate the broader emotional implications of the opposition, and Claude more frequently emphasised existential or philosophical dimensions of the poetic imagery. Importantly, these differences affected only the depth of interpretation rather than the accuracy of contrast recognition itself.

From the perspective of the psycholinguistic framework proposed in the present study, this behaviour indicates that contemporary Large Language Models process explicit contrast primarily through computational operations functionally corresponding to linguistic input processing**,** contextual integration**,** and prediction formation (see Figure 2). Explicit lexical markers substantially constrain the range of possible interpretations, allowing transformer architectures to construct stable contextual representations with relatively little computational uncertainty. The observed consistency across models therefore reflects the availability of strong linguistic evidence rather than similarity to human cognitive processing.

A more informative pattern emerged in the interpretation of the excerpt from Warsan Shire:

No one leaves home unless
home is the mouth of a shark.

All three models successfully recognised the existence of contrast and reconstructed its central conceptual meaning. They consistently interpreted the passage as opposing the conventional image of *home* as a place of safety to its metaphorical transformation into a source of mortal danger. Furthermore, all three models correctly inferred that the poem portrays forced migration not as a voluntary decision but as an act of survival, demonstrating a sophisticated reconstruction of the author's communicative intention.

The principal difference concerned the typological classification of the contrast rather than its interpretation**.** ChatGPT classified the passage as explicit contrast, arguing that both conceptual poles are directly verbalised within the text. Gemini and Claude, by contrast, classified the same passage as implicit contrast, emphasising that the opposition emerges through metaphorical reinterpretation rather than through conventional lexical opposition. Although these classifications differ, they do not affect the interpretative outcome. All three models converged on essentially the same explanation of the poem's central meaning.

This observation is psycholinguistically significant because it demonstrates that typological classification and semantic interpretation represent partially independent processes**.** While the models disagreed regarding the formal classification of the contrastive structure, they nevertheless reconstructed highly similar conceptual and pragmatic interpretations. The findings therefore suggest that successful literary interpretation in transformer-based language models depends more strongly on the quality of contextual representation than on the formal categorisation of contrastive structures.

Table 4. Interpretation of Explicit Contrast by Large Language Models

Poetic excerpt	Model	Contrast detected	Type correctly identified	Conceptual poles	Function of contrast	Author's intention
O. Stomina	ChatGPT	✓	✓	✓	✓	✓
	Gemini	✓	✓	✓	✓	✓
	Claude	✓	✓	✓	✓	✓
L. Kybalka	ChatGPT	✓	✓	✓	✓	✓
	Gemini	✓	✓	✓	✓	✓
	Claude	✓	✓	✓	✓	✓
Warsan Shire	ChatGPT	✓	✓	✓	✓	✓
	Gemini	✓	✗	✓	✓	✓
	Claude	✓	✗	✓	✓	✓

The results obtained for explicit contrast indicate that contemporary Large Language Models exhibit a highly stable strategy of interpretation whenever the linguistic structure of the text provides sufficient evidence for establishing oppositional relations. Variability among the models is limited primarily to differences in explanatory style and typological classification, whereas recognition of contrast, reconstruction of conceptual oppositions, and interpretation of communicative intention remain remarkably consistent. These findings establish an important computational baseline against which the substantially greater challenges posed by implicit contrast can be evaluated in the following section.

Unlike explicit contrast, implicit contrast proved to be a considerably more demanding task for all three Large Language Models. Whereas overtly marked oppositions provided reliable lexical and syntactic cues, hidden contrastive relations required the models to infer semantic opposition from the internal organisation of the poetic image itself. The differences observed in this part of the experiment therefore reveal not only the level of model performance but also the specific interpretative strategies through which transformer-based systems reconstruct implicit meaning.

The most revealing case was the excerpt by A. Voitenko:

А в нього у очах не світять зорі,
а в нього тільки присмак цигарок

All three models failed to identify the implicit contrast in this passage and concluded that the excerpt did not provide sufficient evidence for establishing an oppositional relation. This result is especially important because human participants generally recognised the contrast successfully. The difference suggests that the models were less able than human readers to rely on associative, culturally familiar, and emotionally grounded interpretations of the opposition between elevated romantic imagery and ordinary physical reality.

This failure should not be interpreted simply as a random error. Nor can it be reduced entirely to a failure of the prompt. Since all three models received the same instruction and produced the same cautious response, the result more plausibly reflects the limited amount of explicit semantic evidence available in the excerpt. The passage does not contain strong lexical markers of opposition, and the contrast emerges through culturally conventional associations: *stars in the eyes* evoke inspiration, hope, or romantic elevation, whereas *the taste of cigarettes* evokes bodily fatigue, routine, and material reality. Human readers were able to reconstruct this opposition through shared linguistic intuition and cultural experience. The models, by contrast, treated the passage more conservatively because the surface-level linguistic evidence did not sufficiently constrain the interpretation.

From the perspective of the functional analogue model proposed in the present study, this case illustrates a limitation of contextual representation and inferential reconstruction in LLMs. The models were able to process the literal meanings of the images but did not reorganise them into a contrastive conceptual frame. In terms of Figure 2, the difficulty appears at the level corresponding to frame reconfiguration and inferential interpretation, where human readers can draw on embodied and culturally grounded experience, whereas LLMs rely primarily on distributed textual regularities.

A different pattern emerged in the analysis of O. Kadalov's line:

Я батарейка, що працює навіть з мінусовим зарядом

All three models detected a hidden opposition between exhaustion and continued functioning. ChatGPT and Claude classified the passage as implicit contrast, while Gemini treated it as explicit because the incompatible elements are directly verbalised in the phrase *negative charge*. Despite this disagreement in classification, the models converged in their semantic interpretation. They associated the image with psychological endurance, depletion, resilience, and the ability to continue acting despite the absence of internal resources.

This case demonstrates that LLMs can reconstruct implicit contrast successfully when the text contains sufficiently strong internal semantic cues. The metaphor of a battery provides a familiar conceptual frame, while the phrase *negative charge* introduces a violation of expected functionality. The models were therefore able to identify the contrast by detecting incompatibility within the same conceptual domain. Unlike the Voitenko excerpt, where the contrast depends heavily on cultural and affective associations, Kadalov's image provides a clearer internal structure that can be processed through contextual embeddings and probabilistic semantic association.

The third implicit example, taken from Ocean Vuong, produced the most theoretically significant contrast between human and machine interpretation:

Milkflower petals on the street
like pieces of a girl's dress.

Unlike most human participants, who failed to identify the hidden contrast, all three Large Language Models detected an implicit opposition and classified the passage as contrastive. Their interpretations converged around the opposition between beauty, tenderness, fragility, and implied violence, loss, or destruction. This result might appear to suggest that the models outperformed human readers. However, a closer analysis shows that the nature of their success was fundamentally different from human literary interpretation.

The models did not reconstruct the broader historical and cultural context of *Aubade with Burning City*. Because the experimental instruction restricted them to the excerpt itself, their responses were based on internal semantic relations rather than on external literary or historical

knowledge. They identified a plausible contrast between delicate floral imagery and the implied violence associated with the image of a torn girl's dress, but this interpretation remained generalised. It was not grounded in the specific cultural and historical frame of the poem.

This finding is central to the argument of the present study. LLMs may successfully construct a coherent interpretation of implicit contrast even when they do not access the broader experiential or cultural knowledge required for human literary understanding. Their success results from probabilistic reconstruction of semantic relations within the excerpt rather than from the activation of lived experience, cultural memory, or historically situated interpretation. Human readers, by contrast, often refrained from interpretation when they recognised that the available context was insufficient. The models generated the most probable coherent explanation; the human participants more frequently evaluated the limits of their own interpretative certainty.

Table 5. Interpretation of Implicit Contrast by Large Language Models

Poetic excerpt	Model	Contrast detected	Type correctly identified	Conceptual poles	Function	Author's intention
A. Voitenko	ChatGPT	✗	—	—	—	Partial
	Gemini	✗	—	—	—	Partial
	Claude	✗	—	—	—	Partial
O. Kadalov	ChatGPT	✓	✓	✓	✓	✓
	Gemini	✓	✗	✓	✓	✓
	Claude	✓	✓	✓	✓	✓
Ocean Vuong	ChatGPT	✓	✓	Partial	✓	Partial
	Gemini	✓	✓	Partial	✓	Partial
	Claude	✓	✓	Partial	✓	Partial

The findings indicate that contemporary LLMs do not process implicit contrast uniformly. When the text provides strong internal semantic cues, as in the Kadalov excerpt, the models are capable of reconstructing sophisticated implicit oppositions. When the contrast depends primarily on culturally conventional associations and affective imagery, as in Voitenko's excerpt, the models may adopt an overly cautious strategy and refrain from identifying contrast. When the excerpt contains highly salient internal semantic tension, as in Ocean Vuong's image, the models can generate a coherent contrastive interpretation even without reconstructing the full cultural and historical frame. Thus, their success depends less on human-like understanding than on the density and distribution of semantic cues available within the text.

The comparative analysis demonstrates that the limitations observed in the interpretation of implicit contrast are systematic rather than incidental. Although all three Large Language Models successfully reconstructed many contrastive relationships, their errors followed recurrent patterns that provide valuable insights into the strengths and limitations of transformer-based language processing. Rather than reflecting random inaccuracies, these responses reveal characteristic computational strategies employed when explicit linguistic evidence is insufficient for unambiguous interpretation.

The first type of error concerns the conservative interpretation strategy adopted when the textual evidence supporting a contrastive relationship is relatively weak. This pattern is illustrated by the excerpt from A. Voitenko, where all three models preferred not to identify contrast despite the fact that most human participants reconstructed the intended opposition successfully. Instead of generating a speculative interpretation, the models concluded that the available linguistic evidence did not justify establishing a contrastive relation. Such behaviour suggests that LLMs tend to privilege semantic reliability over interpretative creativity when confidence in the contextual representation is low.

The second type of limitation involves dependence on the density of semantic cues within the text**.** The comparison between the Voitenko and Kadalov excerpts illustrates this phenomenon particularly clearly. Both passages express implicit contrast without overt adversative markers. However, the Kadalov metaphor contains an internally contradictory conceptual structure that provides stronger statistical evidence for reconstructing semantic opposition. Consequently, all three

models successfully interpreted the intended contrast. These findings indicate that LLM performance depends not simply on whether contrast is explicit or implicit but on the degree to which the relevant conceptual relations are recoverable from the linguistic material itself.

A third limitation concerns the reconstruction of culturally and experientially grounded meaning**.** Human readers frequently draw upon autobiographical memory, cultural knowledge, emotional experience, and shared social assumptions when interpreting literary texts. Transformer-based language models have no direct access to such sources of knowledge. Instead, they approximate culturally meaningful interpretations through statistical regularities learned from textual corpora. This distinction becomes particularly important whenever the intended interpretation depends on symbolic associations extending beyond the information explicitly represented in the discourse.

The results also demonstrate that successful responses produced by LLMs should not automatically be interpreted as evidence of human-like understanding. The Ocean Vuong excerpt provides the clearest example. All three models generated coherent interpretations of the implicit contrast, whereas most human participants declined to propose a definitive explanation. Nevertheless, closer examination suggests that the models achieved this result by constructing the most probable semantic hypothesis consistent with the available linguistic evidence rather than by reconstructing the historical, cultural, or experiential background of the poem. Their responses therefore illustrate the effectiveness of probabilistic contextual modelling rather than conceptual understanding in the psycholinguistic sense.

From the perspective of the functional analogue framework, these observations clarify both the similarities and the limits of computational language processing. Contemporary Large Language Models perform a number of interpretative functions that correspond to human language comprehension, including contextual integration, prediction, semantic association, and discourse-level inference. However, these functions operate entirely within a computational architecture based on distributed statistical representations. They do not involve embodied cognition, episodic memory, emotional experience, or culturally situated conceptualisation, all of which play an essential role in human literary interpretation.

Recent research in mechanistic interpretability supports this conclusion. Contemporary transformer models develop increasingly sophisticated contextual representations across multiple processing layers, enabling them to capture long-range semantic dependencies and complex discourse relations. At the same time, these distributed representations remain fundamentally different from the conceptual and experiential knowledge structures that support human cognition. Consequently, similarities between human and LLM interpretations should be understood as examples of functional correspondence rather than evidence of shared cognitive mechanisms.

Large Language Models can successfully reproduce a substantial proportion of the interpretative functions involved in contrast processing, particularly when semantic relationships are strongly represented within the text itself. Their limitations become apparent when successful interpretation requires the integration of implicit cultural knowledge, emotionally grounded conceptual experience, or extensive pragmatic reconstruction. These findings further demonstrate that similarities in interpretative outcome do not imply equivalence of cognitive architecture and therefore should not be interpreted as evidence of human-like language understanding.

Taken together, the human experiment and the computational analysis demonstrate that explicit and implicit contrast differ substantially in their psycholinguistic complexity. While explicit contrast was interpreted successfully by both human participants and Large Language Models, implicit contrast revealed systematic differences in the mechanisms underlying interpretation. These findings provide the empirical basis for the discussion presented in the following section. The principal similarities and differences observed in the experiment are summarised in ***Table 6***.

The comparative findings provide empirical support for the concept of functional analogues developed throughout the present study. Contemporary Large Language Models successfully reproduce a substantial proportion of the interpretative functions involved in contrast processing,

particularly when these functions depend on contextual integration, semantic association, and discourse-level prediction. However, the experiment also demonstrates clear limits to this correspondence. Human interpretation continues to depend on embodied cognition, cultural experience, and pragmatic reasoning, whereas computational systems rely exclusively on representations derived from textual data. Similarity of interpretative outcome therefore should not be interpreted as evidence of cognitive equivalence but rather as the result of different mechanisms performing comparable interpretative functions.

Table 6. Comparative Psycholinguistic Strategies of Human Readers and Large Language Models in Contrast Interpretation

Aspect of interpretation	**Ukrainian readers**	**Large Language Models**
Recognition of explicit contrast	Relies on lexical knowledge and semantic memory	Relies on contextual embeddings and statistical lexical patterns
Interpretation of explicit contrast	Stable and highly accurate	Stable and highly accurate
Recognition of implicit contrast	Depends on contextual reasoning and background knowledge	Depends on the density of semantic cues in the text
Cultural and historical knowledge	Integrated through lived experience and cultural memory	Approximated through statistical regularities in training data
Inferential reasoning	Draws on presuppositions, conceptual knowledge, and pragmatic inference	Based on contextual representations and probabilistic prediction
Response to insufficient context	Interpretation may be suspended because of uncertainty	The most probable coherent interpretation is generated
Principal limitation	Variability resulting from individual knowledge and experience	Difficulty reconstructing culturally grounded implicit meanings

6.3. Functional Analogues between Human Contrast Interpretation and Large Language Models

The experimental findings presented in the preceding sections provide empirical support for the concept of functional analogues proposed in this study. Although human readers and Large Language Models frequently produced similar interpretations of both explicit and, in many cases, implicit contrast, these outcomes were achieved through fundamentally different psycholinguistic and computational mechanisms. The comparison therefore supports functional correspondence rather than cognitive equivalence between human language comprehension and transformer-based language models.

This correspondence is summarised in Table 7, which aligns the principal stages of the proposed psycholinguistic model with their functional computational analogues in Large Language Models.

Table 7. Functional Analogues between Human Contrast Interpretation and Large Language Models

Human psycholinguistic stage	**Functional analogue in Large Language Models**	**Principal computational function**
Linguistic input processing	Tokenisation and contextual embeddings	Initial representation of linguistic input
Prediction formation	Next-token prediction	Generation of probabilistic expectations
Contrast detection	Contextual attention mechanisms	Identification of semantic inconsistency and contrastive cues
Semantic integration	Multi-layer contextual representations	Integration of information across the discourse
Frame reconfiguration	Dynamic updating of contextual representations	Reorganisation of semantic relations during processing
Inferential interpretation	Distributed semantic inference	Reconstruction of implicit conceptual relationships
Pragmatic evaluation	Probabilistic response generation	Selection of the most coherent contextually appropriate interpretation

The correspondence presented in Table 7 should not be interpreted as establishing one-to-one equivalence between psycholinguistic stages and specific components of transformer architectures. Rather, it identifies computational operations that perform broadly comparable interpretative functions while relying on fundamentally different representational principles. This framework provides the conceptual basis for the discussion of similarities and differences between human and computational language processing presented in the following section.

7. Discussion

The findings of the present study support the central proposition that human language comprehension and transformer-based Large Language Models may achieve functionally comparable interpretative outcomes while relying on fundamentally different processing mechanisms. Similarity in interpretation should therefore not be regarded as evidence of cognitive equivalence but rather as the result of different representational systems performing comparable interpretative functions. This distinction constitutes the theoretical foundation of the concept of functional analogues proposed in the present study.

The degree of functional correspondence, however, differs according to the linguistic realisation of contrast. Human participants and all three Large Language Models demonstrated highly consistent interpretations of explicit contrast, indicating that overt linguistic markers effectively constrain the range of possible interpretations. This finding is consistent with contemporary psycholinguistic models, according to which explicit semantic relations activate well-established lexical and conceptual representations during the earliest stages of language comprehension (Kintsch, 1998; Hagoort, 2019). Comparable interpretative outcomes in transformer-based language models appear to arise from contextual representations generated through self-attention and probabilistic prediction rather than from semantic memory in the psycholinguistic sense.

A substantially different pattern emerged in the interpretation of implicit contrast. As explicit linguistic cues decreased, the similarity between human and computational interpretation became progressively weaker. Successful interpretation increasingly depended on contextual integration, conceptual restructuring, inferential reasoning, and pragmatic evaluation. The experimental findings therefore provide empirical support for the proposed seven-stage psycholinguistic model by demonstrating that the greatest variability arises during higher-order stages of processing, particularly frame reconfiguration, inferential interpretation, and pragmatic evaluation.

These findings are consistent with contemporary neurocognitive models of language comprehension, which view meaning construction as a predictive and integrative process rather than the sequential decoding of isolated linguistic units (Kutas & Federmeier, 2011; Hagoort, 2019). The present study extends these models by demonstrating that implicit contrast places particularly high demands on the coordination of contextual prediction, semantic integration, and inferential reasoning. At the same time, it develops previous psycholinguistic research on contrast interpretation (Yuldasheva, 2025) by showing that several mechanisms identified in human language comprehension have computational counterparts that perform broadly comparable interpretative functions in transformer-based language models despite relying on fundamentally different representational principles.

The findings also contribute to the ongoing discussion concerning the nature of language understanding in contemporary Large Language Models. Recent studies have shown that transformer-based architectures achieve remarkably high performance on tasks involving metaphor interpretation, pragmatic inference, and Theory of Mind (Strachan et al., 2024). Nevertheless, the present experiment indicates that successful performance should not automatically be interpreted as evidence of human-like understanding. Throughout the study, similar interpretations frequently emerged through fundamentally different mechanisms: human participants relied on semantic memory, embodied knowledge, cultural experience, and pragmatic reasoning, whereas ChatGPT,

Gemini, and Claude generated context-sensitive interpretations through probabilistic prediction and distributed contextual representations.

This conclusion broadly supports the position advanced by Bender and Koller (2020), who argue that linguistic performance alone should not be equated with semantic understanding. This interpretation is also consistent with recent discussions distinguishing statistical language modelling from human semantic understanding (Mitchell & Krakauer, 2023). At the same time, the present findings suggest that the distinction between human cognition and statistical language modelling is more nuanced than a simple opposition between "understanding" and "prediction". The consistent performance of contemporary Large Language Models demonstrates that statistical learning can reproduce many complex interpretative functions observed in human language processing, even though the underlying cognitive and computational mechanisms remain fundamentally different.

Finally, the results complement recent research in mechanistic interpretability, which seeks to explain how transformer architectures encode semantic information across multiple processing layers (Nanda, 2024; Anthropic, 2025). Whereas such studies primarily investigate the internal organisation of language models, the present research approaches the problem from the opposite direction, beginning with experimentally validated psycholinguistic mechanisms and identifying computational operations that perform comparable interpretative functions. This perspective provides a theoretically grounded framework for comparing human language comprehension and artificial intelligence without assuming either cognitive equivalence or complete incomparability.

Accordingly, the functional analogues proposed in the present study should be interpreted as correspondences between information-processing functions rather than one-to-one mappings between individual psycholinguistic stages and specific transformer layers.

Beyond the interpretation of the experimental findings, the present study makes several theoretical and methodological contributions to both psycholinguistics and Artificial Intelligence research.

From a psycholinguistic perspective, the study demonstrates that contrast interpretation should be understood as a multilevel cognitive process rather than the identification of semantic opposition alone. The experimental findings support a distinction between contrast recognition and contrast interpretation, showing that successful comprehension depends on a sequence of interrelated cognitive operations extending from linguistic decoding and semantic integration to frame reconfiguration, inferential reasoning, and pragmatic evaluation. The results provide empirical support for the proposed seven-stage psycholinguistic model, demonstrating that explicit and implicit contrast engage different levels of cognitive processing. While explicit contrast is generally resolved during the earlier stages of language comprehension, implicit contrast increasingly depends on higher-order interpretative mechanisms. The study therefore extends previous psycholinguistic research on contrast (Yuldasheva, 2025) by explaining how these mechanisms operate during the interpretation of authentic contemporary literary discourse.

The findings also contribute to a more comprehensive understanding of contrast as a psycholinguistic phenomenon. Rather than representing a static semantic relation between opposing concepts, contrast emerges as a dynamic process of meaning construction whose cognitive complexity varies according to the linguistic realisation of the contrastive structure and the degree to which successful interpretation depends on contextual, cultural, and pragmatic knowledge.

From the perspective of Artificial Intelligence, the study proposes an alternative framework for evaluating Large Language Models. Most existing research measures model performance primarily through benchmark accuracy, reasoning tasks, or standard Natural Language Processing evaluations. Although these approaches provide valuable information about model capabilities, they reveal relatively little about the interpretative functions underlying successful language processing. By comparing experimentally observed psycholinguistic mechanisms with computational operations rather than isolated linguistic outputs, the present study offers a more fine-grained framework for analysing the similarities and differences between human and artificial language processing.

The concept of functional analogues constitutes the principal methodological contribution of the study. Instead of asking whether Large Language Models "understand" language in the human sense, the proposed framework examines which interpretative functions they successfully reproduce, under what conditions their performance converges with human interpretation, and where systematic divergence occurs**.** This perspective enables psycholinguistic theory to contribute directly to the evaluation of contemporary Artificial Intelligence systems while preserving the conceptual distinction between computational language processing and human cognition.

More broadly, the proposed framework provides a foundation for future interdisciplinary research integrating psycholinguistics, cognitive linguistics, computational linguistics, and Artificial Intelligence. It illustrates how experimentally validated psycholinguistic models can complement computational evaluations by providing theoretically grounded criteria for investigating increasingly sophisticated language technologies.

The findings of the present study should be interpreted in light of several limitations. Although the experimental corpus was substantially expanded to include thirty authentic twenty-first-century Ukrainian and English-language poetic excerpts**,** the human experiment involved forty native speakers of Ukrainian. Future cross-linguistic studies involving participants from different linguistic and cultural backgrounds are needed to determine the extent to which the interpretation of implicit contrast depends on language-specific and culture-specific cognitive mechanisms. The computational comparison was limited to three contemporary transformer-based Large Language Models (ChatGPT, Gemini, and Claude). While these systems represent some of the most advanced publicly available LLMs, future research should include models based on alternative architectures, multimodal systems, and specialised literary reasoning models in order to evaluate whether the patterns observed in the present study generalise across different AI technologies.

A further limitation concerns the experimental design. To ensure direct comparability between human participants and Large Language Models, all models were evaluated using a single standardised prompt under identical analytical conditions. No follow-up prompts, iterative clarification, or additional system instructions were provided during the evaluation, and the identical prompt reproduced in Appendix A was applied to all thirty poetic excerpts and to all three Large Language Models. Although this approach ensured methodological consistency and experimental reproducibility, it did not allow us to examine the influence of alternative prompting strategies, interactive dialogue, or agent-based reasoning on the interpretation of literary contrast. Future studies should also investigate whether more exploratory system instructions or interactive prompting strategies modify the interpretation of implicit contrast. Moreover, the present investigation relied primarily on behavioural evidence. Future research should therefore integrate behavioural experiments with event-related potentials (ERP), eye-tracking, and functional magnetic resonance imaging (fMRI) to investigate the temporal dynamics of the psycholinguistic stages proposed in the present model. Combining such neurocognitive evidence with recent advances in mechanistic interpretability may provide a more comprehensive account of the functional correspondence between human language comprehension and transformer-based language models, thereby strengthening the interdisciplinary integration of psycholinguistics and Artificial Intelligence research.

8. Conclusions

The present study demonstrates that human readers and contemporary Large Language Models may arrive at comparable interpretations of literary contrast while relying on fundamentally different psycholinguistic and computational mechanisms. The comparison of explicit and implicit contrast shows that the degree of functional correspondence between the two systems depends on the cognitive complexity of the interpretative task. Explicit contrast was interpreted with a high level of agreement because overt linguistic markers provided reliable cues for semantic processing. By contrast, implicit contrast revealed substantially greater divergence, particularly when successful

interpretation depended on contextual integration, cultural knowledge, inferential reasoning, and pragmatic evaluation.

A principal contribution of the study is the development and empirical validation of a seven-stage psycholinguistic model of contrast interpretation**.** The findings demonstrate that explicit and implicit contrast engage different stages of this process to different degrees, confirming the theoretical distinction between contrast recognition and contrast interpretation. While explicit contrast is generally resolved during linguistic processing and semantic integration, implicit contrast increasingly depends on frame reconfiguration, inferential interpretation, and pragmatic evaluation.

The study also proposes a framework of functional analogues that relates human psycholinguistic processes to computational operations in transformer-based LLMs without assuming cognitive or structural equivalence. Rather than comparing isolated linguistic outputs, the proposed framework enables the analysis of interpretative functions performed by humans and LLMs under identical experimental conditions. The findings demonstrate that similarities in interpretative performance are most appropriately understood as instances of functional correspondence rather than evidence of shared cognitive architecture, thereby providing a theoretically grounded basis for comparing human language comprehension and artificial intelligence.

From a methodological perspective, the study introduces a psycholinguistically informed approach to evaluating the interpretative capabilities of LLMs using authentic literary discourse. Unlike conventional benchmark-based evaluations, the proposed methodology focuses on the cognitive functions involved in reconstructing explicit and implicit meaning. It may also be extended to the investigation of other cognitively demanding discourse phenomena, including metaphor, irony, analogy, presupposition, subtext, and indirect communication.

Future research should combine behavioural experiments with neuropsycholinguistic methods, including event-related potentials (ERP), eye-tracking, and functional magnetic resonance imaging (fMRI)**,** in order to investigate the temporal dynamics of contrast interpretation. It should also include multilingual participant groups, broader literary corpora, and LLMs representing different computational architectures. Integrating psycholinguistic experimentation with recent advances in LLM interpretability may contribute to a more precise understanding of the relationship between human language comprehension and computational language processing.

Appendix A. Standard Prompt Used for the LLM Experiment

You are participating in a research study on the interpretation of contrast in contemporary poetry. Analyse the following poetic excerpt using only the text provided. Do not search the Internet and do not use any external sources or background knowledge about the author or the poem. Base your analysis exclusively on the excerpt. If the excerpt does not provide sufficient information to answer a question confidently, explicitly state that the necessary contextual information is missing. Answer each question separately. First, determine whether the text contains contrast and answer only "Yes" or "No". Then identify the exact fragment in which the contrast is realised. Next, determine the type of contrast by choosing only one option: Explicit contrast or Implicit contrast. After that, explain between which concepts, images, or semantic domains the contrast is constructed. Then explain the function of the contrast within the poem. Finally, explain the probable authorial intention based solely on the excerpt without inferring information that is not supported by the text. Be concise, precise, and rely exclusively on textual evidence.

LLM Inference Settings

- Temperature: 0
- Top-p: 1.0
- Max output tokens: default model settings
- One response generated for each excerpt
- No follow-up prompting
- No external retrieval
- Identical prompt applied to all models

References

Anthropic. (2024). *The Claude 3 Model Family: Opus, Sonnet, Haiku*.

Anthropic. (2025). *Tracing the Thoughts of a Large Language Model*.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5185–5198). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.463

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., . . . Liang, P. (2021). *On the opportunities and risks of foundation models*. arXiv. https://doi.org/10.48550/arXiv.2108.07258

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems, 33*, 1877–1901.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long and Short Papers)* (pp. 4171–4186). Association for Computational Linguistics. https://doi.org/10.18653/v1/N19-1423

Fillmore, C. J. (1982). Frame semantics. In *Linguistics in the Morning Calm* (pp. 111–137). Hanshin Publishing.

Gemini Team, Anil, R., Borgeaud, S., Alayrac, J.-B., Yu, J., Soricut, R., Schalkwyk, J., Dai, A. M., Hauth, A., Millican, K., Silver, D., Johnson, M., Antonoglou, I., Schrittwieser, J., Glaese, A., Chen, J., Pitler, E., Lillicrap, T., Lazaridou, A., . . . Dean, J. (2023). *Gemini: A family of highly capable multimodal models*. arXiv. https://doi.org/10.48550/arXiv.2312.11805

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences, 9*(9), 416–423. https://doi.org/10.1016/j.tics.2005.07.004

Hagoort, P. (2019). The neurobiology of language beyond single-word processing. *Science, 366*(6461), 55–58. https://doi.org/10.1126/science.aax0289

He, Y., Zheng, W., Dong, Y., Zhu, Y., Chen, C., & Li, J. (2025). *Towards global-level mechanistic interpretability: A perspective of modular circuits of large language models*. *Proceedings of the 42nd International Conference on Machine Learning (ICML 2025)*, *Proceedings of Machine Learning Research, 267*, 22865–22880. https://proceedings.mlr.press/v267/he25x.html

Heimersheim, S., & Nanda, N. (2024). *How to use and interpret activation patching*. *arXiv*. https://doi.org/10.48550/arXiv.2404.15255

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge University Press.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146*, 23–49. https://doi.org/10.1016/j.brainres.2006.12.063

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press.

Mitchell, M., & Krakauer, D. C. (2023). The debate over understanding in AI's large language models. *Proceedings of the National Academy of Sciences, 120*(13), Article e2215907120. https://doi.org/10.1073/pnas.2215907120

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rogers, A., Kovaleva, O., & Rumshisky, A. (2020). A primer in BERTology: What we know about how BERT works. *Transactions of the Association for Computational Linguistics, 8*, 842–866. https://doi.org/10.1162/tacl_a_00349

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Lawrence Erlbaum Associates.

Strachan, J. W. A., Albergo, D., Borghini, G., Pansardi, O., Scaliti, E., Gupta, S., Saxena, K., Rufo, A., Panzeri, S., Manzi, G., Graziano, M. S. A., & Becchio, C. (2024). Testing theory of mind in large language models and humans. *Nature Human Behaviour, 8*(7), 1285–1295. https://doi.org/10.1038/s41562-024-01882-z

Van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience, 20*(4), 580–591. https://doi.org/10.1162/jocn.2008.20054

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*, 5998–6008.

Yuldasheva, L. P. (2025). Psycholinguistic parameters of the perception and interpretation of linguistic contrast. *Naukovi Zapysky. Seriia: Filolohichni Nauky, 2*(214), 305–313. https://doi.org/10.32782/2522-4077-2025-214.2-38