МРНТИ 16.21.33 DOI: 10.59102/kufil/2026/iss3pp185-199

A. Seisebayev[1], E. Suleimenova[1], A. Groce[2]
[1]Al-Farabi Kazakh National University, Almaty 050040, Republic of Kazakhstan
[2]American Councils for International Education, Washington, D.C. 20036, USA

ENGLISH OXYMORONIC CONSTRUCTIONS IN USE: CORPUS AND NATIVE-SPEAKER EVIDENCE WITH AI-ASSISTED INTERPRETATION (A PILOT STUDY)

This study explores how oxymoronic constructions are interpreted in contemporary English, particularly in cases where frequent usage may weaken the perception of semantic contradiction. The research focuses on fixed expressions with formally antonymic structure and examines the extent to which their paradoxical meaning is preserved or neutralized in actual language use.

The empirical argument draws on two evidence streams: patterns of frequency and contextual stabilization in the Corpus of Contemporary American English (COCA), and native-speaker judgments of contradiction and stylistic markedness. Outputs from ChatGPT and Gemini are used as an auxiliary interpretative benchmark rather than as independent linguistic evidence. This triangulated design permits comparison between usage, speaker perception, and the metalinguistic explanations produced by AI tools.

The findings suggest that formal antonymy alone does not guarantee the perception of contradiction. High-frequency expressions with a terminological function (e.g., civil war, virtual reality) tend to lose their oxymoronic effect, while more expressive constructions (e.g., bittersweet, deafening silence) retain their figurative force. At the same time, AI-generated interpretations appear sensitive to formal structure but less effective in capturing pragmatic nuances.

The study contributes to the understanding of oxymoron as a dynamic and gradable phenomenon rather than a fixed stylistic category. It proposes a model of gradable oxymoronicity that reflects the transition from rhetorical expressiveness to lexicalized usage. The proposed approach may also be applied to the analysis of similar constructions in other languages and discourse contexts.

Key words: oxymoron, semantic contradiction, lexicalization, corpus analysis, artificial intelligence, cognitive interpretation, language models, COCA, native speaker perception, English linguistics.

MAIN PROVISIONS

The present study is based on several theoretical and methodological principles related to the interpretation of oxymoronic constructions in contemporary English. The research is situated within the fields of corpus linguistics, cognitive linguistics, lexical semantics, and pragmatics and examines oxymoron as a dynamic linguistic phenomenon rather than as a purely stylistic device.

The study proceeds from the traditional understanding of oxymoron as a combination of semantically opposite or incompatible elements that produces a paradoxical effect and contributes to expressiveness [1; 2; 3]. Such an interpretation provides the basis for identifying constructions with formal antonymic structure and remains an important point of departure for the analysis. Recent experimental studies have shown that native speakers evaluate oxymoronic expressions not only according to semantic opposition, but also with respect to conventionality, acceptability, familiarity, and interpretability, suggesting that oxymoronicity is influenced by cognitive rather than purely formal semantic factors [14].

At the same time, the research adopts a cognitive perspective according to which oxymoronic constructions may function as meaningful conceptual structures rather than simple logical contradictions. Conceptual blending theory [4] explains how apparently incompatible meanings can be integrated into a unified conceptual space, while conceptual metaphor theory [7] demonstrates

ISSN 2788-7979 (online)

how opposing semantic domains may interact in the process of meaning construction. From this perspective, oxymoron is viewed as a mechanism of conceptual integration rather than merely a rhetorical anomaly.

The study also takes into account pragmatic approaches to meaning interpretation. In accordance with relevance theory and lexical pragmatics [5; 6], the perception of contradiction depends not only on semantic structure but also on discourse context, communicative intention, and conventional patterns of usage. Consequently, formally contradictory expressions may be interpreted by speakers as coherent and semantically acceptable units.

A central assumption of the research concerns the processes of lexicalization and semantic change. Expressions that originally function as stylistically marked oxymorons may gradually lose their paradoxical force through frequent use and become conventionalized lexical units. This tendency is consistent with contemporary phraseological research, which argues that conventional figurative expressions become conventionalized lexical units and may differ in the degree of perceived figurativeness [10]. In some cases, they develop terminological meanings and cease to be perceived as contradictory. This tendency is particularly evident in expressions such as civil war and virtual reality, which retain formal semantic opposition but function as stable lexical items in contemporary English.

Another important principle of the study is the corpus-based approach to linguistic analysis. The research relies on authentic language data from the Corpus of Contemporary American English (COCA) rather than solely on dictionary definitions or linguistic intuition. Corpus evidence makes it possible to examine frequency, contextual variation, genre distribution, and indicators of lexicalization, providing an empirical basis for evaluating the status of oxymoronic constructions in actual language use.

The study further assumes that oxymoronicity is not a binary category. Instead, it should be viewed as a gradable phenomenon whose realization depends on the degree of semantic conflict preserved within an expression. Some constructions continue to function as active oxymorons, while others demonstrate partial neutralization or complete semantic stabilization. This perspective allows oxymoronicity to be analyzed as a continuum rather than as a fixed stylistic category.

An additional methodological principle of the research is the use of artificial intelligence as an analytical instrument. Interpretations generated by large language models are treated as an auxiliary source of metalinguistic analysis, not as evidence of community usage or native-speaker cognition. Consequently, AI outputs cannot override corpus patterns or survey judgments; their role is to show how two models explain the same constructions and to identify possible differences between surface-semantic and pragmatic reasoning. Recent research demonstrates that large language models are increasingly capable of recognizing and interpreting figurative language, including metaphorical and other semantically non-literal constructions. However, their interpretations often rely on surface linguistic features and may not fully capture deeper pragmatic and contextual meanings, especially when figurative interpretation requires pragmatic inference and world knowledge [11; 12; 13].

Finally, the study uses methodological triangulation with differentiated evidential roles. COCA is the primary source for claims about frequency, contextual distribution, and conventionalization, while the native-speaker survey is the primary source for claims about perceived contradiction and stylistic salience. AI-generated interpretations serve only as a secondary comparative instrument. The central argument therefore rests on convergence, qualification, or divergence between corpus usage and speaker judgments, with AI outputs used to illuminate rather than determine the interpretation.

Taken together, these theoretical and methodological principles form the basis of the proposed model of gradable oxymoronicity, according to which oxymoronic constructions may range from strongly paradoxical and stylistically marked expressions to fully lexicalized and semantically neutralized units.

INTRODUCTION

The oxymoron is traditionally defined as a rhetorical figure based on the combination of semantically opposite elements within a single expression, producing a paradoxical effect. Constructions such as *deafening silence* or *sweet sorrow* generate semantic tension, draw attention to the unconventional combination of meanings, and perform an expressive function [1; 2; 3]. However, in contemporary language use, a number of expressions that formally exhibit antonymic structure function as neutral or even terminological units (e.g., *civil war*, *virtual reality*, *original copy*), without evoking a sense of contradiction or stylistic markedness among native speakers. This raises an important question about how such expressions are actually interpreted in contemporary English. Recent corpus-based and experimental research has likewise shown that the interpretation of oxymoronic constructions depends not only on formal semantic opposition but also on contextual factors, conventionalization, and speakers' cognitive judgments, while corpus evidence helps identify recurrent antonymic patterns in authentic language use [14; 15].

Modern cognitive approaches suggest that oxymoron should not be regarded merely as a logical anomaly or a narrow rhetorical device. Instead, it can be viewed as a potentially stable cognitive structure. Theories such as conceptual blending [4], relevance theory [5], and discourse conventions [6] allow antonymic combinations to be interpreted as a form of cognitive integration. In this process, initially contradictory elements are combined into a unified meaning. Within cognitive linguistics, such phenomena may also be interpreted through the lens of conceptual metaphor, where opposing meanings are integrated into a single cognitive model [7]. As these expressions are used more frequently within specific discourse genres, they may gradually lose the sense of semantic conflict between their components. These processes are closely related to lexicalization, whereby originally motivated expressions lose their internal semantic transparency and function as unified nominative units [8].

These observations suggest the need to reconsider the category of oxymoron. While traditionally treated as a binary phenomenon (present vs. absent), contemporary empirical and cognitive approaches allow it to be viewed as a gradable property that varies depending on factors such as frequency, genre, degree of lexicalization, and cognitive accessibility. Under this perspective, expressions previously interpreted as prototypical oxymorons may evolve into idiomatic or even terminological units, losing their paradoxical force.

The present study aims to empirically test this hypothesis using English word combinations commonly identified as oxymoronic in pedagogical, methodological, and reference literature. These include expressions such as *bittersweet*, *open secret*, *clearly confused*, *pretty ugly*, *civil war*, *virtual reality*, *original copy*, *deafening silence*, *minor miracle*, and *seriously funny*. These constructions differ in terms of expressiveness, frequency, and native speaker perception.

It examines how semantic conflict is preserved or weakened in such constructions. It also considers the factors that shape their interpretation in contemporary English. The analysis focuses on fixed word combinations characterized by formal antonymy and considers the extent to which the oxymoronic effect remains explicit in actual usage, as well as the alignment between semantic structure, AI-based interpretation, and native speaker perception.

To achieve this objective, a three-stage methodological framework is employed: (1) corpus analysis of frequency and usage contexts in the Corpus of Contemporary American English (COCA); (2) interpretation of the same expressions using two large language models (ChatGPT and Gemini) based on standardized prompts; and (3) experimental validation through a survey of native English speakers, focusing on perceived contradiction, stylistic markedness, and meaning interpretation.

The methodological foundation of the study is based on the assumption that expressions exhibiting oxymoronic features can be categorized into four provisional types. These include: (1) lexical (classical) oxymorons that retain paradoxical meaning across contexts (e.g., bittersweet, deafening silence); (2) contextual oxymorons in which the conflict emerges only under specific conditions (e.g., pretty ugly, clearly confused); (3) terminological or pseudo-oxymorons that have

ISSN 2788-7979 (online)

lost their internal contradiction due to semantic stabilization (e.g., virtual reality, civil war); and (4) lexicalized idioms in which semantic reinterpretation neutralizes the contradiction (e.g., original copy, minor miracle). This typology allows oxymoronicity to be analyzed as a continuum rather than as a discrete category.

The study is grounded in the assumption that oxymoronicity is not a fixed formal property, but a dynamic characteristic influenced by frequency, lexicalization, and interpretive stability. As a result, expressions that formally preserve antonymic structure may be perceived as neutral or even literal in contemporary usage, particularly when their function has stabilized and their rhetorical force has diminished.

The main contribution of this study lies in triangulating corpus evidence about use with experimental evidence from native-speaker judgments, while using AI-generated interpretations as an auxiliary analytical comparison. Within a unified cognitive-pragmatic framework, this design makes it possible to identify convergence and divergence between formal structure, contextual conventionalization, and perceived contradiction, thereby refining the boundaries of oxymoronicity and supporting a model of its gradual attenuation in language.

MATERIALS AND METHODS

Research Material

The material used in this study includes several types of data and focuses on expressions with formal antonymic structure. The primary linguistic material comprises English word combinations exhibiting features of formal antonymy (e.g., *bittersweet*, *deafening silence*, *open secret*, *civil war*, *virtual reality*, *original copy*, *pretty ugly*, *seriously funny*, *clearly confused*, *minor miracle*).

In addition to the oxymoronic constructions themselves, the research material includes:

1. corpus data obtained from the Corpus of Contemporary American English (COCA) [9] including frequency and contextual usage;

2. survey data collected from native speakers of English, reflecting their intuitive perception of contradiction and stylistic markedness;

3. interpretations of the selected expressions generated by large language models (ChatGPT and Gemini).

The two empirical sources have distinct functions: COCA documents frequency and contextual conventionalization, whereas the survey documents perceived contradiction and stylistic markedness. AI-generated interpretations are included as an auxiliary analytical comparison and are not treated as a third, equivalent source of linguistic evidence.

Principles of Unit Selection

Selection followed four explicitly operationalized steps. The first three determined whether a candidate was eligible for analysis; the fourth guided the composition of a deliberately heterogeneous pilot sample. This procedure is consistent with corpus-based work that treats antonymic relations as a principal criterion for detecting oxymoronic constructions in authentic language data [15].

1. Formal configuration: a candidate had to be a single lexical item or a recurrent two-component expression whose components permit at least one independently interpretable combination of opposed or semantically incompatible meanings.

2. Prior recognition: a candidate had to be attested in English-language stylistic, rhetorical, pedagogical, or reference discussions as an oxymoron or potentially contradictory expression. This requirement prevented the creation of items solely for the experiment.

3. Corpus attestation: a candidate had to yield at least 20 tokens in COCA, permitting inspection of a minimum sample of 20 concordance lines. Items with fewer tokens or only isolated creative occurrences were excluded.

4. Maximum-variation coverage: from the eligible pool, the final set was chosen to span the pre-analytical range from expressive and context-dependent constructions to lexicalized, idiomatic, and terminological units. This criterion guided sampling and did not predetermine the final classification.

ISSN 2788-7979 (online)

Accordingly, the ten expressions constitute a purposive maximum-variation pilot sample rather than a statistically representative sample of English oxymorons. Their preliminary grouping was heuristic and was subsequently checked against corpus contexts and survey responses.

Research Hypothesis

The study tests two related hypotheses through descriptive comparison across the ten expressions.

H1 (variation hypothesis): if formal antonymy were sufficient to produce an oxymoronic effect, all selected expressions would elicit similarly high contradiction judgments. Instead, perceived contradiction is expected to vary substantially across expressions despite their formally contrastive structure.

H2 (conventionalization hypothesis): expressions showing stable terminological or lexicalized use in COCA are expected to receive lower rates of contradiction judgments than expressions whose corpus contexts retain an expressive or figurative function.

For transparent evaluation in this pilot study, a majority threshold of 50% was used for survey judgments. H1 is supported when at least one formally contrastive expression falls below that threshold and the expressions show clear between-item variation. H2 is supported when items with terminological or lexicalized corpus patterns receive lower contradiction rates than active figurative items. The account would not be supported if all ten expressions received majority contradiction judgments and corpus inspection revealed no stable neutral or terminological uses; H2 would also be challenged if conventionalized items were judged at least as contradictory as expressive items.

Because the survey includes 13 participants and the item set is purposively selected, these decision rules are descriptive and hypothesis-generating; they do not support population-level inferential claims. AI outputs are examined separately as an auxiliary comparison and do not determine whether either hypothesis is supported.

Research Questions

In accordance with the stated hypotheses, the following research questions were formulated:

1. Is the perception of semantic conflict preserved in fixed expressions?
2. Does the weakening of the oxymoronic effect correlate with frequency and lexicalization?
3. To what extent do AI-generated interpretations align with native speaker responses?

Research Design and Methods

To address these questions, a triangulated, multi-stage procedure combined corpus analysis with a native-speaker survey and an auxiliary AI-assisted interpretative comparison. The sequence supports reproducibility but does not assign equal evidential weight to its components.

Step 1. Preliminary Analytical Classification

At the initial stage, each expression was assigned a preliminary analytical evaluation in terms of the expected degree of semantic conflict, ranging from active paradox to complete neutralization. This classification was heuristic in nature and served as a baseline for comparison with results obtained at subsequent stages.

Step 2. Corpus Analysis (COCA)

To determine frequency and contextual dynamics, the Corpus of Contemporary American English (COCA) was used. For each expression, the following parameters were recorded:

– absolute frequency (at the time of analysis);
– typical usage contexts (based on a sample of 20–30 instances);
– genre distribution (fiction, spoken, academic, news);
– indicators of lexicalization (structural stability, reduced figurativeness, terminological usage).

Additionally, the analysis considered semantic features such as:

1. the presence of active stylistic contrast;
2. evidence of neutralization through frequent usage;
3. terminological status;
4. semantic reinterpretation of components (e.g., the use of *pretty* as an intensifier in *pretty ugly*).

The results were systematized in tabular form and supported by illustrative corpus examples.

ISSN 2788-7979 (online)

Step 3. AI-Assisted Interpretative Comparison

Each expression was interpreted using two large language models – ChatGPT and Gemini – based on a standardized prompt with the following structure:

Expression: [X]

1. Explain the meaning of the expression "[X]" in contemporary English.
2. Indicate whether the combination contains a semantic contradiction between its components.
3. State whether the expression is used primarily as a figurative device or as a conventionalized expression in modern usage.
4. Provide a brief explanation.

The interpretations were compared according to the following parameters:

– identification or rejection of semantic contradiction;
– recognition of lexicalization;
– pragmatic status (figurative vs. neutral);
– type of argumentation (semantic, contextual, pragmatic).

Full responses are provided in Appendix A.

Step 4. Native Speaker Survey

At the final stage, a survey of native speakers of American English was conducted (n = 13). Participants were presented with the same expressions without any indication of their stylistic classification. For each expression, the following questions were asked:

1. How do you understand this expression?
2. Does this expression sound contradictory to you?
3. How does this expression feel to you?

Data analysis combined quantitative and qualitative approaches. Quantitative analysis involved calculating response distributions, while qualitative analysis focused on identifying typical interpretative strategies. Special attention was paid to the distinction between logical recognition of contradiction and intuitive stylistic perception. Because this is a pilot study, percentages are interpreted descriptively; the 50% threshold indicates whether a majority of respondents perceived contradiction and is not an inferential significance test.

Evidence Roles and Disagreement Protocol

The central claim is carried by the two empirical sources, each matched to the question it can answer. COCA has priority for statements about frequency, contextual distribution, lexicalization, and terminological stabilization. The survey has priority for statements about whether speakers perceive contradiction and stylistic markedness. ChatGPT and Gemini outputs provide a controlled comparison of metalinguistic explanations but are not treated as observations of language use or human cognition.

Disagreements were resolved by matching the claim to the relevant source rather than by averaging unlike measures. Corpus evidence determines usage status, and survey evidence determines perceptual status. If those sources diverge, the item is classified as mixed or intermediate and both findings are reported. AI output never overrides either empirical source; disagreement with AI is analyzed as a limitation or difference in interpretative strategy.

Summary of Methodological Framework

The study is structured as a sequential, triangulated research procedure involving the following stages:

1. preliminary classification of expressions;
2. corpus analysis;
3. auxiliary AI-assisted interpretative comparison;
4. experimental validation (survey);
5. comparative analysis of results.

This sequence enables a transparent comparison of formal-semantic structure, attested usage, and speaker perception while keeping AI interpretation in an explicitly auxiliary role.

ISSN 2788-7979 (online)

RESULTS

The results are organized around two empirical evidence streams: COCA patterns of use and native-speaker survey judgments. Outputs from ChatGPT and Gemini are reported as an auxiliary interpretative comparison. This organization makes it possible to compare formal structure and contextual conventionalization with perceived contradiction without treating AI output as equivalent to linguistic behavior.

Corpus Analysis (COCA)

The corpus analysis revealed significant differences in both frequency and contextual dynamics of the selected expressions. The most frequent items were *civil war* (20,749 occurrences), *bittersweet* (2,238), and *virtual reality* (2,046), indicating a high degree of lexicalization or terminological stabilization. Less frequent expressions, such as *clearly confused* (25) and *original copy* (35), nevertheless demonstrate stability within certain professional and conversational contexts.

The analysis of usage contexts allowed the identification of three dominant tendencies:

- **Terminologization** (*civil war*, *virtual reality*): expressions function as neutral designations, with semantic conflict fully neutralized.
- **Lexicalization and grammaticalization** (*open secret*, *pretty ugly*, *seriously funny*): components lose their original oppositional meaning and undergo pragmatic shift.
- **Retention of expressiveness** (*bittersweet*, *deafening silence*, *minor miracle*): contexts activate paradoxical meaning, and expressions retain their figurative force.

AI-Assisted Interpretative Comparison

Interpretations generated by ChatGPT and Gemini confirmed the presence of formal antonymy in most cases. However, differences between the models were observed:

- ChatGPT tends to account for pragmatic context, genre specificity, and lexicalization patterns;
- Gemini more often emphasizes the literal combination of components, identifying oxymoronicity at a structural level.

For example, both models classified *civil war* and *virtual reality* as fully lexicalized or terminological expressions with neutralized contradiction. In contrast, *bittersweet* and *deafening silence* were interpreted as figurative constructions retaining rhetorical force.

The expression *clearly confused* is particularly noteworthy: both models identified it as semantically neutral, indicating that the modifier *clearly* does not reinforce contradiction but instead resolves it pragmatically.

Full interpretations are provided in Appendix A.

Native Speaker Perception (n = 13)

The survey of native speakers of American English revealed a notable discrepancy between formal structure and perceived contradiction. Only three expressions were identified as contradictory by more than 30% of respondents:

- *open secret* – 69.2%
- *deafening silence* – 46.2%
- *bittersweet* – 30.8%

All other expressions, including *civil war* (7.7%), *virtual reality* (15.4%), and *clearly confused* (0%), were largely not perceived as paradoxical.

These results indicate cognitive adaptation to constructions that may have originally conveyed semantic conflict but are now processed as neutral expressions due to conventionalization.

Qualitative responses further showed that respondents distinguish between **logical contradiction** and **stylistic perception**: some participants acknowledged a contradiction on a logical level but did not experience it as stylistically salient.

Triangulated Comparative Analysis

To compare the two empirical evidence streams and place the auxiliary AI interpretations alongside them, a comparative table was constructed, including:

- corpus frequency (COCA);

ISSN 2788-7979 (online)

– AI-assisted interpretations (ChatGPT and Gemini; auxiliary comparison);
– percentage of respondents perceiving contradiction;
– stylistic perception (e.g., irony, expressiveness, neutrality);
– summarized interpretations provided by respondents.

The comparison was interpreted according to the pre-specified evidence roles: corpus contexts determined claims about conventionalization, survey responses determined claims about perceived contradiction, and AI outputs were used only to compare explanatory strategies. Where corpus and survey evidence diverged, the construction was retained as a mixed or intermediate case rather than forced into a single consensus category.

Table 1 – Comparative Overview of COCA Evidence, AI-Assisted Interpretation, and Native-Speaker Judgments

Expression	COCA Frequency	AI Interpretation	Perceived Contradiction (%)	Stylistic Perception	Respondent Interpretation
bittersweet	2238	figurative, contradiction preserved	30.8	emotional, poetic	mixture of joy and sadness
deafening silence	128	figurative, contradiction preserved	46.2	paradoxical, intense	“loud silence”
open secret	442	partially neutralized	69.2	ironic	widely known “secret”
civil war	20749	terminologized	7.7	neutral	internal conflict
virtual reality	2046	terminologized	15.4	technical	simulated/digital reality
original copy	35	partially neutralized	15.4	formal	original document
pretty ugly	97	weakened contradiction	23.1	colloquial, ironic	fairly ugly
seriously funny	151	weakened contradiction	23.1	colloquial	very funny
clearly confused	25	neutralized	0	neutral	obviously confused
minor miracle	42	figurative	38.5	expressive	small miracle

DISCUSSION

The results obtained in this study allow for a refinement of the understanding of oxymoron as a linguistic phenomenon in contemporary language use. Contrary to the traditional view, which treats oxymoron as a static stylistic device based on logical contradiction, the findings demonstrate that the presence of formal antonymy within an expression does not necessarily generate a perception of paradox, either in usage or in cognitive interpretation.

A key finding of the study is the consistent discrepancy between componential semantics and cognitively and pragmatically driven interpretation. Expressions such as *civil war* and *virtual reality*, despite their underlying antonymic structure, function as terminologized units with high frequency and genre stability. They are perceived by native speakers as unified lexical items that do not evoke rhetorical tension. This is supported both by corpus data and by survey results.

ISSN 2788-7979 (online)

The evidence did not converge for every item. Open secret is conventionalized in corpus usage, yet 69.2% of respondents perceived a contradiction, the highest rate in the survey. Following the disagreement protocol, it is therefore classified as a conventionalized but perceptually salient expression rather than as fully neutralized. This counterexample qualifies H2 by showing that conventionalization can weaken, but does not invariably erase, oxymoronic awareness.

These findings are consistent with recent empirical research on oxymoronicity. Bolognesi et al. (2024) demonstrated that native speakers evaluate oxymoronic expressions not solely on the basis of semantic opposition, but also according to factors such as familiarity, conventionality, acceptability, and interpretability. Similarly, the present study shows that highly conventionalized expressions, despite their formally antonymic structure, are frequently interpreted as semantically coherent lexical units rather than as stylistically marked oxymorons [14].

At the same time, expressions such as *bittersweet*, *deafening silence*, and *minor miracle* demonstrate a stable preservation of figurative expressiveness. Respondents consistently identify these expressions as emotionally or stylistically marked, and both AI models classify them as retaining an active paradoxical effect. This suggests that oxymoronicity persists in contexts where expressions remain pragmatically salient and are associated with expressive or evaluative discourse.

Particular attention should be given to the semantic shift observed in constructions such as *pretty ugly*, *seriously funny*, and *clearly confused*. In these cases, one of the components – typically a modifier – loses its original semantic value and functions instead as an intensifier or pragmatic marker. This process of grammaticalization leads to the weakening or elimination of the initial contradiction, resulting in idiomatic expressions with reduced stylistic markedness.

The interpretations provided by ChatGPT and Gemini further highlight differences in processing strategies. While Gemini tends to identify oxymoronicity based on formal semantic opposition, ChatGPT incorporates pragmatic context and usage patterns into its analysis. This distinction reflects a broader contrast between formal semantic processing and cognitively oriented, discourse-based interpretation.

Native speaker perception, in turn, appears to be guided primarily by pragmatic function rather than formal structure. Expressions that are frequent and contextually stabilized are not perceived as oxymoronic, even when they retain an underlying semantic contradiction. This indicates that oxymoronicity is not an inherent property of the expression itself, but rather an emergent effect dependent on usage, context, and cognitive processing.

The comparison of corpus evidence and native-speaker judgments, supplemented by AI-assisted interpretation, supports a provisional model of gradable oxymoronicity. Unlike the traditional binary approach, which classifies expressions as either oxymoronic or non-oxymoronic, the present pilot data indicate a continuum along which expressions exhibit varying degrees of semantic conflict.

Within this model, three general levels can be distinguished:

1. Fully active oxymorons, which preserve paradoxical semantics and are perceived as figurative (*bittersweet*, *deafening silence*);

2. Partially weakened oxymorons, where semantic shift or pragmatic neutralization reduces the contradiction (*pretty ugly*, *seriously funny*, *minor miracle*);

3. Neutralized or lexicalized expressions, functioning as conventional or terminological units (*civil war*, *virtual reality*, *clearly confused*).

Thus, oxymoron should be understood as a dynamic linguistic phenomenon, sensitive to frequency, pragmatic function, and discourse context, and capable of evolving from an expressive rhetorical construction into a neutral lexical unit.

CONCLUSION

The aim of the present study was to determine the extent to which semantic conflict is preserved or weakened in English expressions with formal antonymic structure and to identify the factors influencing their interpretation in contemporary usage. To achieve this aim, corpus evidence was

ISSN 2788-7979 (online)

compared with native-speaker judgments, while AI-generated interpretations were used as an auxiliary analytical benchmark.

The descriptive results support H1 and provide partial support for H2. Perceived contradiction varied substantially across expressions, demonstrating that formal antonymy does not necessarily produce a shared perception of paradox. Terminological and strongly lexicalized items generally received low contradiction rates, although open secret shows that conventionalization can coexist with persistent awareness of contradiction. Oxymoronicity is therefore best treated, provisionally, as a gradable and dynamic characteristic shaped by usage, lexicalization, pragmatic function, and genre-specific stabilization.

In some cases, complete terminologization is observed, as in *civil war* and *virtual reality*, where expressions function as neutral lexical units. In other cases, semantic shift and grammaticalization affect one of the components, as seen in *pretty ugly*, *seriously funny*, and *clearly confused*, leading to the attenuation or elimination of the original contradiction. Only a limited number of expressions, such as *bittersweet* and *deafening silence*, retain their expressive and rhetorical force and are perceived as prototypical oxymorons.

The AI-assisted comparison shows that language-model outputs are sensitive to formal antonymic structure but vary in their treatment of contextual and pragmatic factors. These outputs remain diagnostic rather than evidential: claims about usage derive from COCA, and claims about perception derive from the survey.

The scientific novelty of the study lies in triangulating corpus evidence and native-speaker judgments within a cognitive-pragmatic framework while using AI-assisted interpretation as a controlled auxiliary comparison. The resulting provisional model of gradable oxymoronicity reflects varying degrees of semantic conflict in fixed expressions and conceptualizes oxymoron as a dynamic continuum rather than a discrete category.

The proposed approach can be extended to the analysis of other languages and discourse types, including cross-linguistic and digital linguistic studies. Future research may focus on:

– expanding the corpus of analyzed oxymoronic expressions;
– incorporating additional national and specialized corpora;
– integrating sociolinguistic and cross-cultural parameters;
– developing automated classification models based on corpus data and artificial intelligence methods.

APPENDIX A

Examples of AI-Generated Interpretations of Oxymoronic Constructions:

The following examples illustrate interpretations generated by ChatGPT and Gemini using a standardized prompt format. The models were asked to explain the meaning of each expression, identify the presence of semantic contradiction, determine whether the expression functions as a figurative device or a conventionalized expression, and provide a brief explanation.

Expression: bittersweet

ChatGPT:

The expression bittersweet refers to an experience combining both positive and negative emotions, such as happiness mixed with sadness. The components bitter and sweet create an oxymoronic structure due to their semantic opposition. However, in contemporary English the expression functions primarily as a conventionalized lexicalized unit expressing emotional ambivalence.

Gemini:

Bittersweet describes experiences that are simultaneously pleasant and painful. Although structurally oxymoronic, the expression has become highly conventionalized in modern English and is commonly interpreted as a unified emotional concept rather than as a stylistically marked contradiction.

Expression: deafening silence

ISSN 2788-7979 (online)

ChatGPT:

Deafening silence refers to a silence that feels intense, overwhelming, or emotionally significant. The expression contains a clear semantic contradiction because deafening normally describes extremely loud sound, whereas silence denotes the absence of sound. The phrase functions primarily as a figurative stylistic device.

Gemini:

The expression describes a silence so noticeable that it feels overwhelming. The contradiction between deafening and silence creates a classic oxymoron. Despite frequent usage, the expression still retains expressive rhetorical force.

Expression: civil war

ChatGPT:

Civil war denotes an armed conflict within a single country. Although the phrase may appear oxymoronic because civil is associated with civic order and war with violence, the contradiction is no longer perceived in contemporary English. The expression functions as a fully lexicalized terminological unit.

Gemini:

The phrase civil war is primarily used as a conventionalized political and historical term. While the expression may appear contradictory when civil is interpreted as “polite” or “peaceful,” in actual usage it refers to conflict between citizens and is not generally perceived as oxymoronic.

Expression: virtual reality

ChatGPT:

Virtual reality refers to an immersive computer-generated environment. The combination formally suggests a contradiction between virtual and reality, but in modern usage the phrase functions as a fully conventionalized technical term with neutralized semantic conflict.

Gemini:

Although virtual reality combines apparently opposing concepts, the expression has become a standard technological term. Speakers typically interpret it literally as a digital simulated environment rather than as a rhetorical oxymoron.

Expression: clearly confused

ChatGPT:

The phrase clearly confused means that someone’s confusion is obvious to an observer. Although the expression may initially appear contradictory, clearly modifies the observer’s perception rather than the state of confusion itself. Therefore, the phrase is not interpreted as a true oxymoron.

Gemini:

The expression combines clearly and confused, which may appear contradictory at the surface level. However, in actual usage, the phrase functions as a conventional descriptive expression in which clearly acts as an intensifier indicating evident confusion.

Due to space limitations, Appendix A presents only representative examples of AI-generated interpretations. The complete dataset included all analyzed expressions.

INFORMATION ON THE USE OF ARTIFICIAL INTELLIGENCE

Artificial intelligence tools (ChatGPT and Gemini) were used as auxiliary instruments during the analytical and language-editing stages of the study. All analytical interpretations, conclusions, and final revisions were independently verified by the authors.

REFERENCES

1 Leech G. N. A linguistic guide to English poetry (London, Longman, 1969).
2 Baldick C. The Oxford dictionary of literary terms, 3rd ed. (Oxford, Oxford University Press, 2008).

3 Arutyunova N. D. Logika i poetika oksyumorona [The logic and poetics of oxymoron] (Moscow, Nauka, 1990). [in Russian]
4 Fauconnier G., Turner M. The way we think: Conceptual blending and the mind's hidden complexities (New York, Basic Books, 2002).
5 Giora R. On our mind: Salience, context, and figurative language (Oxford, Oxford University Press, 2003).
6 Zufferey S. Lexical pragmatics and theory of mind: The acquisition of connectives (Amsterdam, John Benjamins, 2010).
7 Lakoff G., Johnson M. Metaphors we live by (Chicago, University of Chicago Press, 1980).
8 Moon R. Fixed expressions and idioms in English: A corpus-based approach (Oxford, Clarendon Press, 1998).
9 Davies M. The corpus of contemporary American English (COCA). Available at: https://www.english-corpora.org/coca/ (accessed 12.05.2026).
10 Dobrovol'skij D. Phraseology and Figurative Language: Some Basic Concepts and Future Prospects // Yearbook of Phraseology. – 2025. – Vol. 16(1). – P. 5–50. DOI: 10.1515/phras-2025-0003.
11 Ichien N., Stamenković D. and Holyoak K. J. (2024), Large Language Model Displays Emergent Ability to Interpret Novel Literary Metaphors, Metaphor and Symbol, 39(4), pp. 296–309.
12 Sanchez-Bayona E. and Agerri R. (2025), Metaphor and Large Language Models: When Surface Features Matter More than Deep Understanding, Findings of the Association for Computational Linguistics: ACL 2025, pp. 17462–17477.
13 Chakrabarty T., Ghosh D., Poliak A., Muresan S. Figurative Language in Recognizing Textual Entailment // Findings of ACL-IJCNLP 2021. – 2021. – P. 3354–3361. DOI: 10.18653/v1/2021.findings-acl.297.
14 Bolognesi M.M., Combei C.R., La Pietra M., Masini F. What Makes an Awfully Good Oxymoron? Language and Cognition. 2024. Vol. 16, No. 1. P. 242–262. https://doi.org/10.1017/langcog.2023.68
15 La Pietra M., Masini F. Oxymorons: A Preliminary Corpus Investigation // Proceedings of the Second Workshop on Figurative Language Processing. – Association for Computational Linguistics, 2020. – P. 176–185. https://doi.org/10.18653/v1/2020.figlang-1.24.

Received: 29.06.2026

Ағылшын тіліндегі оксюморондық конструкциялардың қолданысы: корпустық деректер, тіл иелерінің бағалаулары және жасанды интеллект көмегімен жасалған интерпретация (пилоттық зерттеу)

А.Ж. Сейсебаев[1], Э.Д. Сулейменова[1], А.М. Гросс[2]
[1]Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы 050040, Қазақстан
[2]Халықаралық білім беру жөніндегі американдық кеңестер, Вашингтон қ., Колумбия округі 20036, АҚШ

Бұл мақалада қазіргі ағылшын тіліндегі оксюморондық конструкциялардың интерпретациясы қарастырылады, әсіресе олардың жиі қолданылуы семантикалық қарама-қайшылықты қабылдаудың әлсіреуіне әсер ететін жағдайларға назар аударылады. Зерттеу формалды антонимиялық құрылымы бар тұрақты сөз тіркестеріне бағытталған және олардың парадокстық мағынасының нақты тілдік қолданыста сақталу немесе бейтараптану дәрежесін анықтауды мақсат етеді.

Талдау екі эмпирикалық дерек көзіне сүйенеді: қолданыс жиілігі мен контекстік тұрақтануды көрсететін Corpus of Contemporary American English (COCA) корпусының деректері және қарама-қайшылық пен стильдік айқындықты қабылдауды көрсететін ағылшын тілі иелерінің сауалнама жауаптары. ChatGPT пен Gemini интерпретациялары дербес лингвистикалық дәлел ретінде емес, көмекші талдау құралы ретінде пайдаланылады.

ISSN 2788-7979 (online)

Мұндай триангуляциялық дизайн нақты қолданысты, тіл иелерінің бағалауларын және модельдердің метатілдік түсіндірмелерін салыстыруға мүмкіндік береді.

Зерттеу нәтижелері формалды антонимияның өздігінен семантикалық қарама-қайшылықтың қабылдануын қамтамасыз етпейтінін көрсетеді. Терминологиялық қызмет атқаратын жоғары жиілікті тіркестер (мысалы, civil war, virtual reality) оксюморондық әсерін жоғалтуға бейім болса, экспрессивті құрылымдар (bittersweet, deafening silence) өздерінің бейнелілік қасиетін сақтайды. Сонымен қатар, жасанды интеллект арқылы жасалған интерпретациялар формалды құрылымға сезімтал болғанымен, прагматикалық реңктерді толық жеткізе алмайды.

Зерттеу оксюмороны тұрақты стилистикалық категория ретінде емес, динамикалық әрі градуалды құбылыс ретінде түсіндіруге үлес қосады. Мақалада риторикалық экспрессивтіліктен лексикаланған қолданысқа өту үдерісін сипаттайтын градуалды оксюморондық модель ұсынылады. Ұсынылған тәсілді басқа тілдер мен дискурс түрлеріндегі ұқсас конструкцияларды талдауда да қолдануға болады.

Кілт сөздер: оксюморон, семантикалық қарама-қайшылық, лексикалану, корпустық талдау, жасанды интеллект, тілдік модельдер, когнитивтік интерпретация, COCA, сауалнама, ағылшын тілі.

ӘДЕБИЕТТЕР ТІЗІМІ

1 Лич Г. Н. Ағылшын поэзиясының лингвистикалық талдауы. – Лондон: Longman, 1969.
2 Baldick C. The Oxford dictionary of literary terms. – 3rd ed. – Oxford: Oxford University Press, 2008.
3 Арутюнова Н.Д. Оксюморонның логикасы мен поэтикасы. – Мәскеу: Наука, 1990.
4 Fauconnier G., Turner M. The way we think: Conceptual blending and the mind's hidden complexities. – New York: Basic Books, 2002.
5 Giora R. On our mind: Salience, context, and figurative language. – Oxford: Oxford University Press, 2003.
6 Zufferey S. Lexical pragmatics and theory of mind: The acquisition of connectives. – Amsterdam: John Benjamins, 2010.
7 Lakoff G., Johnson M. Metaphors we live by. – Chicago: University of Chicago Press, 1980.
8 Moon R. Fixed expressions and idioms in English: A corpus-based approach. – Oxford: Clarendon Press, 1998.
9 Davies M. The corpus of contemporary American English (COCA) [Электрондық ресурс]. – URL: https://www.english-corpora.org/coca/ (қаралған күні: 12.05.2026).
10 Dobrovol'skij D. Phraseology and Figurative Language: Some Basic Concepts and Future Prospects // Yearbook of Phraseology. – 2025. – Vol. 16(1). – P. 5–50. DOI: 10.1515/phras-2025-0003.
11 Ichien N., Stamenković D., Holyoak K. J. Large Language Model Displays Emergent Ability to Interpret Novel Literary Metaphors // Metaphor and Symbol. – 2024. – Vol. 39(4). – P. 296–309.
12 Sanchez-Bayona E., Agerri R. Metaphor and Large Language Models: When Surface Features Matter More than Deep Understanding // Findings of the Association for Computational Linguistics: ACL 2025. – 2025. – P. 17462–17477.
13 Chakrabarty T., Ghosh D., Poliak A., Muresan S. Figurative Language in Recognizing Textual Entailment // Findings of ACL-IJCNLP 2021. – 2021. – P. 3354–3361. DOI: 10.18653/v1/2021.findings-acl.297.
14 Bolognesi M.M., Combei C.R., La Pietra M., Masini F. What Makes an Awfully Good Oxymoron? // Language and Cognition. – 2024. – Vol. 16(1). – P. 242–262. DOI: 10.1017/langcog.2023.68.
15 La Pietra M., Masini F. Oxymorons: A Preliminary Corpus Investigation // Proceedings of the Second Workshop on Figurative Language Processing. – Association for Computational Linguistics, 2020. – P. 176–185. DOI: 10.18653/v1/2020.figlang-1.24.

ISSN 2788-7979 (online)
Материал 29.06.2026 баспаға түсті

Английские оксюморонные конструкции в употреблении: корпусные данные и оценки носителей языка с ИИ-ассистированной интерпретацией (пилотное исследование)

А.Ж. Сейсебаев[1], Э.Д. Сулейменова[1], А.М. Гросс[2]
[1]Казахский национальный университет имени аль-Фараби, Алматы 050040, Казахстан
[2]Американские советы по международному образованию , Вашингтон, округ Колумбия 20036, США

В данной статье рассматривается интерпретация оксюморонных конструкций в современном английском языке, особенно в тех случаях, когда высокая частотность употребления способствует ослаблению восприятия семантического противоречия. Исследование сосредоточено на устойчивых словосочетаниях с формально антонимической структурой и направлено на выявление степени сохранения либо нейтрализации их парадоксального значения в реальном языковом употреблении.

Анализ опирается на два эмпирических источника: данные Corpus of Contemporary American English (COCA), отражающие частотность и контекстуальную стабилизацию выражений, и результаты опроса носителей английского языка, отражающие восприятие противоречия и стилистической маркированности. Интерпретации ChatGPT и Gemini используются как вспомогательный аналитический инструмент, а не как самостоятельное лингвистическое доказательство. Такой триангуляционный дизайн позволяет сопоставить реальное употребление, оценки носителей языка и метаязыковые объяснения моделей.

Результаты исследования показывают, что формальная антонимия сама по себе не гарантирует восприятия противоречия. Высокочастотные выражения, выполняющие терминологическую функцию (например, civil war, virtual reality), склонны утрачивать оксюморонный эффект, тогда как более экспрессивные конструкции (bittersweet, deafening silence) сохраняют образную выразительность. Одновременно интерпретации, сгенерированные искусственным интеллектом, демонстрируют чувствительность к формальной структуре, однако менее эффективно отражают прагматические нюансы.

Исследование способствует пониманию оксюморона как динамического и градуального явления, а не фиксированной стилистической категории. В работе предлагается модель градуальной оксюморонности, отражающая переход от риторической экспрессивности к лексикализованному употреблению. Предлагаемый подход может быть применён и к анализу аналогичных конструкций в других языках и типах дискурса.

Ключевые слова: оксюморон, семантическое противоречие, лексикализация, корпусный анализ, искусственный интеллект, языковые модели, когнитивная интерпретация, COCA, опрос, английский язык.

СПИСОК ЛИТЕРАТУРЫ

1 Лич Г. Н. Лингвистический анализ английской поэзии. – Лондон: Longman, 1969.
2 Baldick C. The Oxford dictionary of literary terms. – 3rd ed. – Oxford: Oxford University Press, 2008.
3 Арутюнова Н.Д. Логика и поэтика оксюморона. – М.: Наука, 1990.
4 Fauconnier G., Turner M. The way we think: Conceptual blending and the mind’s hidden complexities. – New York: Basic Books, 2002.
5 Giora R. On our mind: Salience, context, and figurative language. – Oxford: Oxford University Press, 2003.
6 Zufferey S. Lexical pragmatics and theory of mind: The acquisition of connectives. – Amsterdam: John Benjamins, 2010.
7 Lakoff G., Johnson M. Metaphors we live by. – Chicago: University of Chicago Press, 1980.

ISSN 2788-7979 (online)
8 Moon R. Fixed expressions and idioms in English: A corpus-based approach. – Oxford: Clarendon Press, 1998.
9 Davies M. The corpus of contemporary American English (COCA) [Электронный ресурс]. – URL: https://www.english-corpora.org/coca/ (дата обращения: 12.05.2026).
10 Dobrovol'skij D. Phraseology and Figurative Language: Some Basic Concepts and Future Prospects // Yearbook of Phraseology. – 2025. – Vol. 16(1). – P. 5–50. DOI: 10.1515/phras-2025-0003.
11 Ichien N., Stamenković D., Holyoak K. J. Large Language Model Displays Emergent Ability to Interpret Novel Literary Metaphors // Metaphor and Symbol. – 2024. – Vol. 39(4). – P. 296–309.
12 Sanchez-Bayona E., Agerri R. Metaphor and Large Language Models: When Surface Features Matter More than Deep Understanding // Findings of the Association for Computational Linguistics: ACL 2025. – 2025. – P. 17462–17477.
13 Chakrabarty T., Ghosh D., Poliak A., Muresan S. Figurative Language in Recognizing Textual Entailment // Findings of ACL-IJCNLP 2021. – 2021. – P. 3354–3361. DOI: 10.18653/v1/2021.findings-acl.297.
14 Bolognesi M.M., Combei C.R., La Pietra M., Masini F. What Makes an Awfully Good Oxymoron? // Language and Cognition. – 2024. – Vol. 16(1). – P. 242–262. DOI: 10.1017/langcog.2023.68.
15 La Pietra M., Masini F. Oxymorons: A Preliminary Corpus Investigation // Proceedings of the Second Workshop on Figurative Language Processing. – Association for Computational Linguistics, 2020. – P. 176–185. DOI: 10.18653/v1/2020.figlang-1.24.

Материал поступил в редакцию журнала 29.06.2026

ҒТАМР 16.21.39 DOI: 10.59102/kufil/2026/iss3pp199-213

А.А. Солтанбекова[1], Т.Б. Рамазанов[1]*, С.К. Жукенова[2]
[1]Ахмет Байтұрсынұлы атындағы Тіл білімі институты, Алматы, 050010, Қазақстан
[2]Шоқан Уәлиханов атындағы Көкшетау университеті, Көкшетау, 020000, Қазақстан

ҚАЗАҚ ТІЛІНДЕГІ ҚҰРАНДЫ ЕТІСТІКТЕР: ҚҰРЫЛЫМДЫҚ МОДЕЛЬДЕРІ, МАҒЫНАЛЫҚ ТҰТАСТЫҒЫ ЖӘНЕ КЕСТЕЛІК ЖҮЙЕЛЕНУІ

Мақала қазақ тіл біліміндегі күрделі етістіктер жүйесінің құрамында ерекше орын алатын құранды етістіктердің құрылымдық, семантикалық және функционалдық ерекшеліктерін жүйелеуге арналған. Қазақ грамматикасында күрделі етістіктер жан-жақты зерттелгенімен, құранды етістіктердің сөзжасамдық табиғатын, олардың еркін етістікті тіркестерден, идиомаланған және аналитикалық құрылымдардан айырмашылығын кестелік негізде біріздендіріп көрсету қажеттілігі сақталып отыр.

Зерттеудің мақсаты – құранды етістіктерге қатысты ғылыми көзқарастарды жинақтап, олардың құрылымдық модельдерін, мағыналық тұтастығын және сөзжасамдық қызметін жүйелі түрде сипаттау. Негізгі бағыт – есім сөз + көмекші етістік үлгісі арқылы жасалатын құранды етістіктердің жаңа қимыл атауын қалыптастыру қабілетін айқындау, сондай-ақ оларды аналитикалық және идиомаланған етістіктермен салыстырмалы түрде талдау.

Жұмыстың ғылыми маңызы – құранды етістіктерді күрделі етістік жүйесіндегі дербес сөзжасамдық модель ретінде нақтылауында. Практикалық маңызы – құранды етістіктердің құрылымдық және семантикалық ерекшеліктерін кестелік форматта ұсыну арқылы оларды оқыту мен лексикографиялық сипаттауда қолдануға мүмкіндік беруінде.

Зерттеу барысында салыстырмалы-талдау, құрылымдық сипаттау және кестелік модельдеу әдістері қолданылды. Құранды етістіктердің мағыналық тұтастығы,