Automatic detection of Russia-Ukraine war euphemisms

Iryna Dilai[1,*,†], Maksym Davydov[2,†], Anna Feldman[3,†], Olha Oleksyn[1,†], Svitlana Kohut[1,†] and Olha Baranovska[1,†]

[1] *Ivan Franko National University of Lviv, Universytetska Street 1, 79000, Lviv, Ukraine*
[2] *Ukrainian Catholic University, Kozelnytska Street 2, 79026, Lviv, Ukraine*
[3] *Montclair State University, Normal Avenue 1, 07043-1624, Montclair, NJ, USA*

Abstract

Automatic detection of figurative language is one of the major directions in modern NLP. Euphemisms are words or phrases used to mitigate the expression. By and large, they are socially and culturally determined, naming the sensitive entities in an indirect, softened way. The problems of the automatic detection of euphemisms arise when words can be used both literally (non-euphemistically) and euphemistically. We refer to such usages as PETs (potentially euphemistic terms). The attempts to detect/disambiguate euphemisms cross-linguistically have reported a high performance of transformer-based neural models. Nonetheless, such models have not been tested on Ukrainian datasets. The purpose of this endeavor is to test LLMs on the collected, annotated, and processed Ukrainian dataset, exemplified in this paper by the newly coined during the Russia-Ukraine war PETs. Employing prompt engineering, the study has revealed a high performance of GPT-4o and GPT-4o-mini on the Ukrainian PET dataset.

Keywords

Euphemism, automatic detection, NLP, FLP, LLM, prompt engineering, Russia-Ukraine war

1. Introduction

Euphemisms are a linguistic device used to soften or neutralize language that may otherwise be harsh or awkward to state directly. By acting as alternative words or phrases, euphemisms are used daily to maintain politeness, mitigate discomfort, or conceal the truth. As part of war vocabulary, they are used to address sensitive issues, such as death, losses, military attacks, alleviating the mental perception of the harsh reality and often promoting victorious discourse. Euphemistic expressions replace direct references and are characterized by changing the word sentiment towards neutral or more positive. Nonetheless, in the course of time, their positive or neutral sentiment is becoming obliterated; they undergo the "euphemism treadmill" via changing or losing their euphemistic meanings (Pinker, 2003) [1].

Russia-Ukraine war gave rise and prominence to a number of euphemisms in the Ukrainian language: *бавовна* (cotton – "explosion"), *пташка* (birdie – "drone"), *задвохсотити* ("to kill"), *на щиті* ("dead"), etc. The majority of them are relatively new coinages and have not been sufficiently recorded or studied thus far. They are mostly the product of productive semantic derivation patterns accompanied by vivid associations. Focusing on this category of euphemisms in terms of their automatic detection appears to be highly topical and has a clear-cut practical value.

MoMLeT-2025: 7th International Workshop on Modern Machine Learning Technologies, June, 14, 2025, Lviv-Shatsk, Ukraine

* Corresponding author.
† These authors contributed equally.
iryna.dilay@lnu.edu.ua (I. Dilai); maks.davydov@ucu.edu.ua (M. Davydov); feldmana@mail.montclair.edu (A. Feldman); olha.oleksyn@lnu.edu.ua (O. Oleksyn); svitlana.kohut@lnu.edu.ua (S. Kohut); olha.baranovska@lnu.edu.ua (O. Baranovska)
0000-0001-9626-290X (I. Dilai); 0000-0001-7479-8690 (M. Davydov); 0000-0003-4146-5990 (A. Feldman); 0009-0009-3994-1709 (O. Oleksyn); 0000-0002-2570-1415 (S. Kohut); 0000-0002-7723-7458 (O. Baranovska)

Despite being an important element of language use, the figurative nature of euphemisms poses a challenge for natural language processing (NLP). Due to the polysemous nature of the potentially euphemistic terms (PETs), the detection and recognition of their euphemistic usages requires the elaboration of viable mechanisms of word sense disambiguation. The semantic annotation scheme applied to PETs poses difficulties as it needs to consider subtle context-sensitive instances with various shades of meaning. Thus, we hypothesize that drawing on manual annotation of multiple instances of PETs allows for approaching the full specification principle (Lakoff) [2] in the description of the word meaning (conceptual category), which can further feed large language models (LLMs) and train them to detect and recognize euphemistic expressions of the Ukrainian language.

Thus, the aim of this paper is to discover efficient techniques for the automatic detection of Ukrainian euphemisms related to the topic of the Russia-Ukraine war, which presupposes the completion of the following tasks:

- To collect a dataset of Ukrainian war-related euphemisms based on the corpus of modern web communication.
- To elaborate, standardize, and apply the annotation scheme for PETs.
- To elicit key difficulties in the recognition and annotation of the PETs.
- To leverage machine learning techniques for the automatic detection of euphemisms.
- To assess the performance of the models.

2. Related Works

In recent years, there has been a surge of interest in computational approaches to euphemism detection in the NLP community. [3] introduce the recognition of euphemisms and dysphemisms using NLP, generating near-synonym phrases for sensitive topics. [4] propose euphemism detection and identification tasks using masked language modeling with BERT. [5] create an extensive corpus of potentially euphemistic terms (PETs). In [6], they develop a linguistically driven approach for identifying PETs using distributional similarities. BERT-based systems that participated in a shared task on euphemism disambiguation they organized showed promise [7]. [8] experiment with classifying PETs unseen during training. In [9], they perform transformer-based euphemism disambiguation experiments, exploring vagueness as one of the properties of euphemisms.

The work of R. Choenni et al. [10] explores the multilingual and cross-lingual transfer capabilities of LLMs. They find that multilingual LLMs rely on data from multiple languages to a large extent, learning both complementary and reinforcing information. The authors of [11] find cases where transfer learning from out-of-language data in a particular domain performed better than the same-language data in a different domain.

While euphemisms are culturally dependent, the need to discuss sensitive topics in a non-offensive way is universal, suggesting similarities in the way euphemisms are used across languages and cultures. Euphemisms are found across the world’s languages, making them a universal linguistic phenomenon. As such, euphemistic data may have useful properties for computational tasks across languages. A. Feldman and her team have explored this premise by training a multilingual transformer model (XLM-RoBERTa) to disambiguate potentially euphemistic terms (PETs) in multilingual and cross-lingual settings. They have conducted experiments on English, Spanish, Chinese, Turkish, and some other low-resource languages. In line with current trends, they demonstrate that zero-shot learning across languages takes place. They also showcase where multilingual models perform better on the task compared to monolingual models by a statistically significant margin, indicating that multilingual data presents additional opportunities for models to learn about cross-lingual, computational properties of euphemisms. In a follow-up analysis, they focus on universal euphemistic "categories" such as death and bodily functions, among others. They test to see whether cross-lingual data of the same domain is more important than within-language data of other domains to further understand the nature of the cross-lingual transfer.

In June 2024, a special FigLang (Figurative Language Processing) workshop was held in Mexico, where the findings of the shared tasks on multilingual euphemism detection were presented. Among others, [12] tried to test whether Chat GPT can detect euphemisms across multiple languages.

To the best of our knowledge, no similar study on the automatic detection of euphemisms has been conducted in Ukraine or on Ukrainian language material. Nonetheless, there are works with valuable observations related to the linguistic aspect of the newly coined Ukrainian military vocabulary and euphemisms in particular [13–16]. A revised approach to labeling sensitive language related to the ongoing war was proposed in [17].

3. Methods and Materials

The technical solution to the problem of automated Ukrainian euphemisms detection involves prompt engineering for LLMs. We test both zero-shot prompting, which does not contain any examples or demonstrations while interacting with the model, and few-shot prompting, accompanied by illustrations and enabling in-context learning of the model.

The experimental research is based on GPT-4o, the flagship LLM by OpenAI, GPT-4o-mini, and DeepSeek, though other models have also been tested but showed worse performance. GPT-4.5-preview was rejected due to its high pricing at the time of testing. Older models (o1, o3-mini, o1-mini) performed notably worse on smaller datasets and were also rejected. DeepSeek-Chat was chosen as a widely advertised, cheaper alternative to OpenAI models.

We draw on the F1 score to elaborate on class-wise performance of the LLMs. The overall workflow consists of the following stages:

- PET dataset collection.
- PET dataset annotation.
- PET dataset processing.
- Testing zero-shot and few-shot prompting performance of the LLMs.
- Prompt engineering to enhance the performance of the models.

PET samples have been collected from the Polish Automatic Web corpus of the Ukrainian language (PAWUK) [18]. It was built and is being maintained by our partner institution the Linguistic Engineering Group of the Institute of Computer Science of the Polish Academy of Sciences. The corpus contains Ukrainian texts selected from the web pages and social network posts starting from February 24, 2022, and is daily updated. As of March 2025, it consists of over 700 million tokens. The PAWUK is lemmatized and accompanied by automatic POS tagging.

The seed list for the dataset encompasses 21 PETs together with their derivatives featuring the war-related vocabulary. Since most of the identified PETs are polysemous items (e.g., *пташка, бавовна, ціль, мінусувати*, etc.), not always used euphemistically, the key problem both for human annotators and for the AI lies in disambiguating their senses. It can be tackled by accomplishing their fine-grained annotation and elaborating a machine learning model that would achieve high performance in the recognition of euphemistic usages.

The initial stage of testing a euphemism detection model is the collection of a PET dataset. Our dataset consists of 4,258 instances of Ukrainian PETs referring to the ongoing Russia-Ukraine war, encompassing both euphemistic and non-euphemistic usages. Table 1 features the resultant dataset.

When collecting the dataset based on the PAWUK, we were guided by the following principles: (1) tried to get a balanced representation of the PET across the three-year period (2022–2025), (2) tried to represent all wordforms for both euphemistic and non-euphemistic usages, (3) tried to use a corpus-driven approach, proportionately representing euphemistic and non-euphemistic usages, their wordforms, etc., (4) tried to identify and include cases hard to classify: instances of pun, symbolism, intentional vagueness, etc.

The dataset collection consists of the sentences with the PET, ideally, with the preceding and/or following sentences to introduce broader context. The node, e.g., <бавовна>, was enclosed in angle brackets in each sample to facilitate further processing. The annotation stage lies in labeling PETs as euphemistic with a label (0) and non-euphemistic with a label (1).

Table 1
PET dataset description

PET	PET translation (euphemistic)	Euphemistic instances	Non-euphemistic instances	Total number
бавовна	explosion	749	251	1000
(за)бавовнитися	to explode	7	0	7
пташка	drone, aircraft	124	77	201
дискотека	military action	166	334	500
двохсотий ((за)двохсотити)	dead (to kill)	180	20	
трьохсотий (за)трьохсотити	wounded to wound	175	25	200
на щиті	perished	40	10	50
мінусувати	to liquidate	142	58	200
відпрацювати	to shell	46	103	149
мопед	drone	155	95	250
приліт (прилетіти)	hit (to hit)	132	154	286
втомитися	to collapse	119	87	206
ціль	target	101	99	200
м'ясо	cannon fodder	92	108	200
спеціальна воєнна операція	special military operation	50	0	50
зоряні війни	star wars	12	12	25
приземлити	to shoot down	104	46	150
на концерт Кобзона	to die	201	0	201
закобзонити	to kill	3	0	3
батальйон Монако	affluent draft dodgers	10	0	10
за руски/ім кораблем	after the russian warship (expletive)	50	0	50
дружній вогонь	friendly fire	20	0	20
нуль	frontline	151	49	200
	Total	***2,829***	***1,429***	***4,258***

The annotation was done manually by four annotators who are expert linguists. The inter-annotator agreement was measured using *Cohen's kappa (κ).* For the resultant dataset, $\kappa = 0.89$. The annotators were asked to mark the cases of uncertainty, attaching the most likely label to the respective sentences.

The cases of uncertainty encompassed the samples of distinct play on words (pun), in which the euphemistic usage keeps traces of the literal one and cannot be discerned without it, e.g., *зацвіла бавовна.* Also, we identified the cases of the literal use of this PET with a noticeable shade of the new euphemistic sense. In these cases, the PET *бавовна* is used in the sense of "a flower/plant" referred to as a symbol of "(victorious) explosion". Such nuances contribute to the complexity of PET annotation.

The dataset processing pinpoints its major quantitative characteristics. Though we tried to obtain a balanced and representative dataset of the PET, it is limited to (1) the time span (2022–2025) and (2) the web communication register. Thus, the obvious bias will be towards euphemistic usage, often accompanied by metaphor, irony, and sarcasm.

4. Experiment

The performance of LLMs is highly affected by a prompt that is passed to interact with the model and perform the detection of euphemisms. To reduce the mutual impact of data samples on each other, we rejected batching several data samples into one prompt, although batching reduces the overall pricing of data processing without a high impact on the performance.

The experiments were planned to understand the impact of the prompt on LLM performance. Four types of prompts were chosen (Table 2) for experiments with different scopes of additional information provided. The language of the prompts (Ukrainian/ English) did not substantially affect the performance.

Table 2
Prompts used in the experiments

Prompt type	Prompt
Context-free prompt	[Prompt 1] For each sentence in the set, determine whether the term enclosed in angle brackets is used as a euphemism (1) or not (0).
Prompt with a labeled example	[Prompt 2] For each sentence in the set, determine whether the term enclosed in angle brackets is used as a euphemism (1) or not (0). Consider the example of labeling.
Prompt with a dictionary definition	[Prompt 3] You are a linguist. For each sentence in the set, determine whether the term enclosed in angle brackets is used as a euphemism (1) or not (0). Consider the terms to be euphemistic in the context of war, the dictionary definitions are attached. (The dictionary definitions generated by the GPT-4o model are provided in Appendix A).
Prompt with a word list	[Prompt 4] You are a linguist. For each sentence in the set, determine whether the term enclosed in angle brackets is used as a euphemism (1) or not (0). Consider the terms to be euphemistic in the context of war; the list of euphemisms is attached. (The attached list of euphemisms without definitions is provided in Appendix B).

The initial (context-free/zero-shot) prompt was: "For each sentence in the set, determine whether the term enclosed in angle brackets is used as a euphemism (1) or not (0)". Table 3 shows the sample of PET labeling in comparison with the annotators' labeling.

The agreement between the annotators and GPT-4o was estimated. For the PET *бавовна*, the F1 score is equal to 0.77 (Precision = 0.82, Recall = 0.72). For the whole dataset, the F1 score amounts to 0.81, which is rather high, though individual PETs show different performances (from 0.5 to 0.9).

The next step was to check if the performance could be augmented after refining the prompt, providing the AI with a few-shot prompting.

Table 3
Sample PET labeling by expert annotators vs GPT-4o

#	Context	Annotators' label	GPT-4o label	Agreement
1	Сонцезахисні штори не обов'язково вибирати серед темних відтінків. Світлі теж ефективні. Погано із захистом від сонця справляються льон і <бавовна>.	0	0	1
2	Повідомляється, що горять резервуари із паливом. Даних про постраждалих не надходило. За словами жителів, перед початком пожежі чулася потужна <бавовна>, а заграва від вогню зараз видно з різних точок міста.	1	1	1
3	Російським новинам заборонили казати вибух. Вони казали хлопок. А хтось переклав на українську хлопок-<бавовна>.	0	1	0
4	Курськ знову горить. <Бавовна> знову винна.	1	0	0
5	У нас зброя значиться на кожній людині. Є зброя, яка належала <двохсотим> і трьохсотим.	1	0	0
6	Ворожі <прильоти> у Краматорську: пошкоджено залізничне полотно.	1	1	1
7	Рації передали, що за контрольно-пропускним пунктом «<трьохсотий>».	1	1	1
8	Воїни Збройних сил України приземлили чергову російську <пташку> під Бахмутом на Донеччині.	1	0	0

5. Results

The performance of LLMs on the PETs dataset largely depends on the type of the model and the prompt. The DeepSeek-chat model was not significantly affected by the prompt type and its performance was considerably worse than GPT-4o-mini and GPT-4o models. GPT-4o-mini performed unexpectedly better than GPT-4o on context-free prompts, regardless of its smaller size.

Providing definitions of the war-related euphemisms was beneficial for the GPT-4o-mini and GPT-4o models, but the performance boost was considerably higher for the GPT-4o model (+11%).

One of the hypotheses was that the LLM improves performance by utilizing a list of euphemisms in the context of the war without focusing on the meaning of the euphemisms themselves. To prove or disprove this, we provided a list of euphemisms without an explanation of their meaning (Prompt 4). The result turned out to be worse than when providing no word samples at all, as in Prompt 1. The inclusion of 10 random labeled examples of euphemistic and non-euphemistic usage of words in the prompt (Prompt 2) had no significant impact on the performance of the models. Moreover, it significantly increases the prompt size, thereby raising inference costs.

Table 4 shows the performance of all the models tested on Prompts 1, 3, and 4. The results of Prompt 2 are omitted because they do not differ significantly from the results of the context-free prompt.

Table 4
Russia-Ukraine war euphemism detection performance of LLMs with different prompts

Model	Prompt type	Precision	Recall	F1
deepSeek-chat	Context-free prompt	0.699	0.837	0.762
deepSeek-chat	Prompt with a dictionary definition	0.699	0.818	0.754
deepSeek-chat	Prompt with a word list	0.693	0.854	0.765
gpt-4o-mini	Context-free prompt	0.804	0.841	0.822
gpt-4o-mini	Prompt with a dictionary definition	**0.833**	**0.880**	**0.856**
gpt-4o-mini	Prompt with a word list	0.730	0.779	0.754
gpt-4o	Context-free prompt	**0.903**	0.731	0.809
gpt-4o	Prompt with a dictionary definition	0.875	**0.964**	**0.918**
gpt-4o	Prompt with a word list	0.699	0.814	0.752

The detection rate is unevenly distributed among the euphemisms (Table 5). The LLMs' performance on the PETs *відпрацювати, зоряні війни, дискотека, ціль, втомитися* was much worse than the performance on other terms.

Table 5
The euphemism detection performance of the GPT-4o model based on a prompt with a dictionary definition (Prompt 3)

PET	Number of samples	Precision	Recall	F1
бавовна	1007	0.960	0.972	0.967
двохсотий	200	0.942	0.994	0.968
приліт	150	0.969	1	0.984
трьохсотий	200	0.940	0.989	0.964
прилетіти	350	0.878	0.960	0.917
втомитися	206	0.723	0.832	0.773
пташка	201	0.873	0.951	0.911
ціль	200	0.201	0.964	0.333
спеціальна воєнна операція	50	1	0.96	0.980
на щиті	50	0.905	0.974	0.938
приземлити	150	0.893	0.971	0.930
мопед	250	0.910	0.981	0.944
Батальон Монако	10	1	1	1
дискотека	500	0.548	0.968	0.670
зоряні війни	25	0.647	0.917	0.759
за рускім кораблем	50	1	1	1
дружній вогонь	22	1	1	1
на концерт до Кобзона	204	1	1	1
мінусувати	200	0.838	0.992	0.909
м'ясо	200	0.929	0.989	0.958
відпрацювати	149	0.303	0.964	0.461
нуль	200	0.947	0.953	0.950

6. Discussion

The quality of annotation has largely predetermined the performance of the model. Among the major challenges for the annotators were:

1. Labeling instances with play on words (pun).
2. Handling the symbolic usage and metaphoric extensions of different types.
3. Adopting a vantage point, as the same PETs appeared to be euphemistic and have more positive sentiment when referred to enemy losses but looked dysphemistic and acquired negative sentiment when referred to one's own losses (comp., *бавовна в Терноп*олі).
4. Annotators' inner bias.
5. Already noticed euphemism treadmill resulting in the tendency to gradually treat them as rather dysphemistic within broader contexts.

Engineering LLMs' prompts that can best detect euphemistic usages in context involved experimenting with zero-shot and few-shot modes. The highest F1 scores have been achieved by GPT-4o and GPT-4o-mini for the whole dataset while interacting with a prompt accompanied by dictionary definitions of the euphemisms under scrutiny.

It is worth mentioning that the PET dataset is not homogeneous; it comprises clear-cut instances of euphemisms always labeled with (1), which are generally easier to detect, and ambiguous instances of polysemous PETs where either euphemistic or non-euphemistic usages prevail based on the corpus data. The task was also complicated by insufficiency of context in some cases. Besides, some PETs refer to more than one euphemistic category and, as a result, were ignored due to the focus of the prompts on the war-related vocabulary.

Another observation is that though the overall performance of GPT-4o and GPT-4o-mini achieved on the Ukrainian PET dataset is strikingly high, the models often fail to explain why a certain word or phrase is euphemistic (they provide wrong synonyms, hypernyms, or definitions). It demonstrates that even though the correct label has been attached, the models' understanding of the sense/usage is incorrect.

7. Conclusions

The euphemism treadmill illustrates how language evolves in response to societal attitudes and how efforts to soften language often fall short of removing the negative associations that these terms might evoke. It highlights a tension between the desire to use language to be more sensitive and inclusive and the reality that such efforts can sometimes inadvertently create new stigmas.

The study has proven that the war-related euphemisms manifest the vast creative potential of the users and are particularly context-sensitive. As mostly newly coined, euphemisms are a challenging problem for detection and proper understanding by humans, let alone AI. Nonetheless, neural network models relying on efficient techniques can easily recognize them and use them in other applications, including generative AI.

The implications of this research go beyond computational linguistics and NLP. Ukrainian war-related euphemisms designating sensitive topics are a rapidly developing category in the Ukrainian language, reflecting the new reality and its perception. Thus, the results can be of interest to the social sciences.

The prospects of further study lie in testing the models on a larger dataset of Ukrainian PETs belonging to other categories and employing other, more advanced LLMs.

Acknowledgements

This study is supported by the STCU and is part of the broader international collaboration within the IMPRESS-U project #7132 "DARE: Detecting and Recognizing Euphemisms", which is a supplement to the NSF grant #2226006.

Declaration on Generative AI

During the preparation of this work, the authors used GPT-4o in order to: Paraphrase and reword in Prompt 3 (Appendix A) and integrate the generated definitions in the experiment.

References

[1] S. Pinker, The Blank Slate: The Modern Denial of Human Nature. Viking, 2003.

[2] G. Lakoff, Women, Fire, and Dangerous Things. What Categories Reveal about the Mind, University of Chicago Press, Chicago, 1987.

[3] C. Felt, E. Riloff, Recognizing euphemisms and dysphemisms using sentiment analysis, in: Proceedings of the Second Workshop on Figurative Language Processing, pp. 136–145, Online, July 2020, Association for Computational Linguistics. doi:10.18653/v1/2020.figlang-1.20

[4] W. Zhu, H. Gong, R. Bansal, Z. Weinberg, N. Christin, G. Fanti, S. Bhat, Self-supervised euphemism detection and identification for content moderation, in: 42nd IEEE Symposium on Security & Privacy, 2021, arXiv preprint arXiv:2103.16808. doi:10.48550/arXiv.2103.16808

[5] M. Gavidia, P. Lee, A. Feldman, J. Peng. CATs are fuzzy PETs: A corpus and analysis of potentially euphemistic terms, in: Proceedings of the Thirteenth Language Resources and Evaluation Conference, pp. 2658–2671, Marseille, France, June 2022. European Language Resources Association.

[6] P. Lee, M. Gavidia, A. Feldman, J. Peng, Searching for PETs: Using distributional and sentiment-based methods to find potentially euphemistic terms, in: Proceedings of the Second Workshop on Understanding Implicit and Underspecified Language, pp. 22–32, Seattle, USA, July 2022. Association for Computational Linguistics. doi:10.18653/v1/2022.unimplicit-1.4

[7] P. Lee, A. Feldman, J. Peng, A report on the euphemisms detection shared task, in: Proceedings of the 3rd Workshop on Figurative Language Processing (FLP), pp. 184–190, Abu Dhabi, United Arab Emirates (Hybrid), December 2022. Association for Computational Linguistics. doi:10.18653/v1/2022.flp-1.27

[8] S. S. Keh, Exploring euphemism detection in few-shot and zero-shot settings, in: Proceedings of the 3rd Workshop on Figurative Language Processing (FLP), pp. 167–172, Abu Dhabi, United Arab Emirates (Hybrid), December 2022. Association for Computational Linguistics. doi:10.18653/v1/2022.flp-1.24

[9] P. Lee, I. Shode, A. Chirino Trujillo, Y. Zhao, O. E. Ojo, D. Cuevas Plancarte, A. Feldman, J. Peng, FEED PETs: Further Experimentation and Expansion on the Disambiguation of Potentially Euphemistic Terms, in: Proceedings of the 12th Joint Conference on Lexical and Computational Semantics (*SEM 2023), pp. 437–448, Toronto, Canada, July 2023. Association for Computational Linguistics. doi:10.18653/v1/2023.starsem1.38

[10] R. Choenni, D. Garrette, E. Shutova. How do languages influence each other? Studying cross-lingual data sharing during LLM fine-tuning, in: Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, pp. 13244–13257, Singapore, December 2023. Association for Computational Linguistics. doi:10.18653/v1/2023.emnlp-main.818

[11] I. Shode, D. Ifeoluwa Adelani, J. Peng, A. Feldman, Nollysenti: Leveraging transfer learning and machine translation for Nigerian movie sentiment classification, in: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), May 2023. Association for Computational Linguistics. doi: 10.48550/arXiv.2305.10971

[12] T. Firsich, A. Rios, Can GPT-4 Detect Euphemisms across Multiple Languages? in: Proceedings of the 4th Workshop on Figurative Language Processing (FigLang), June 21, 2024, pp. 65–72. Association for Computational Linguistics. doi: 10.18653/v1/2024.figlang-1.9

[13] O. Levchenko, Specifics of Ukrainian military discourse, in: Ucrainica X, Vydala Univerzita Palackého v Olomouci, Olomouc, Czechia, 2023, pp. 199–207.

[14] V. Balazh, Pragmalinguistic aspects of the research into euphemisms and dysphemisms (based on Ukrainian news Telegram channels), New Philology, 89 (2023) 35–42. doi:10.26661/2414-1135-2023-89-5
[15] O. Kharchenko, Ukrainian metaphorical euphemisms during the Russian-Ukrainian war, Printing Horizon, National Technical University of Ukraine "Kyiv Polytechnic Institute named after Igor Sikorsky", Kyiv, Ukraine, 2/14 (2023) 90–101. doi:10.20535/2522-1078.2023.2(14).295247
[16] O. Taranenko, Euphemization in the Ukrainian media discourse of the hybrid war period, Social communication: theory and practice, 4 (2017) 19–27.
[17] D. Stetsenko, When a Language Question Is at Stake. A Revisited Approach to Label Sensitive Content, arXiv:2311.10514v1 [cs.CL], November 17, 2023. URL: https://arxiv.org/abs/2311.10514
[18] W. Kieraś, Ł. Kobyliński, D. Komosińska, B. Nitoń, M. Rudolf, M. Shvedova, A. Zwierzchowska, PAWUK: Polish Automatic Web corpus of UKrainian language, Instytut Podstaw Informatyki PAN, Warszawa 2023. URL: https://pawuk.ipipan.waw.pl.

Appendices

Appendix A. Definitions of euphemisms for Prompt 3

Бавовна – іронічне позначення вибухів, яке виникло через цензуру в російських медіа. Замінює слово «вибух» у контексті ударів по ворожих об’єктах.

Двохсотий – військовий термін, що означає загиблого солдата (походить від кодової назви «вантаж 200» для транспортування тіл загиблих).

Приліт – потрапляння ракети, снаряду або дрону в ціль, зазвичай супроводжується вибухом.

Трьохсотий – військовий термін, що означає пораненого солдата (походить від кодової назви «вантаж 300» для евакуації поранених).

Прилетіти – отримати влучання ракетою чи снарядом, зазвичай використовується щодо обстрілів міст, військових об’єктів або техніки.

Втомитися – евфемізм, яким часто описують стан російських систем ППО або техніки після удару ЗСУ.

Пташка – безпілотник або літальний апарат, який виконує розвідувальні чи ударні завдання.

Ціль – об’єкт, по якому планується завдати удару (наприклад, військова техніка, командний пункт, склад боєприпасів).

Спеціальна воєнна операція – евфемістичний термін, який росія використовує для позначення свого повномасштабного вторгнення в Україну з метою уникнення слова «війна».

Приземлити – збити ворожий літак, безпілотник чи ракету.

Мопед – іронічна назва іранського дрона-камікадзе «Shahed», який використовується для ударів по українській інфраструктурі (через характерний звук двигуна, схожий на мопед).

Батальйон Монако – саркастичний термін для українських багатіїв та політиків, які втекли за кордон під час війни, особливо в дорогі курортні місця на кшталт Монако.

Дискотека – масований обстріл або бомбардування, часто супроводжується вибухами та загравою.

Зоряні війни – протиповітряний бій із застосуванням ППО, коли в небі видно сліди від збитих ракет або дронів.

За рускім кораблем – скорочена форма українського військового мему «Русскій корабль, іді *!», що став символом спротиву російській агресії.

Дружній вогонь – випадковий обстріл своїх військ або техніки, часто через погану координацію або паніку.

На концерт до Кобзона – евфемізм, який означає загибель російських військових чи командирів (Йосип Кобзон – радянський співак, що підтримував російську агресію, помер у 2018 році).

Мінусувати – знищувати ворожу техніку або живу силу (наприклад, «мінуснули танк» – знищили танк).

М’ясо – мобілізовані солдати, яких російське командування кидає в бій без належної підготовки та забезпечення (також відоме як «м’ясні штурми»).

Відпрацювати – завдати удару по ворожій позиції або техніці (наприклад, «артилерія відпрацювала по складу БК»).

Нуль – передова лінія фронту, найнебезпечніше місце, де тривають активні бойові дії.

На щиті – вираз, що означає загибель військового у бою. Походить із давньої традиції, коли загиблих воїнів приносили з поля бою на щитах. У сучасному контексті використовується як синонім терміна «двохсотий».

Appendix B. The list of euphemisms for Prompt 4

Бавовна, двохсотий, приліт, трьохсотий, прилетіти, втомитися, пташка, ціль, спеціальна воєнна операція, на щиті, приземлити, мопед, Батальон Монако, дискотека, зоряні війни, за рускім кораблем, дружній вогонь, на концерт до Кобзона, мінусувати, м’ясо, відпрацювати, нуль.