

HUMANITIES

DOI <https://doi.org/10.51647/kelm.2026.3.12>

METODY LINGWISTYKI KOMBINATORYCZNEJ W OPISIE SŁOWNICTWA EMOCJONALNEGO: ASPEKTY KORPUSOWO-STATYSTYCZNE I KOGNITYWNO-SEMANTYCZNE

Nataliia Bober

*kandydat nauk filologicznych, docent,
docent na katedrze filologii germańskiej,*

*Wydziału Filologii Romańsko-Germańskiej Uniwersytetu Stołecznego im. Borisa Grinchenka w Kijowie
(Kijów, Ukraina)*

ORCID ID: 0000-0002-9639-0562

n.bober@kubg.edu.ua

Adnotacja. Artykuł rozważa metody lingwistyki kombinatorycznej jako narzędzie opisu słownictwa emocjonalnego. Analiza koncentruje się na kolokacji, walencji, dystrybucji, powiązaniach kolostrukuralnych, preferencjach semantycznych, prozodii semantycznej oraz korpusowych środkach asocjacyjnych. Leksem emocjonalny traktowany jest nie jako izolowana nominacja emocji, lecz jako całość, której znaczenie jest wyjaśniane poprzez powtarzające się kombinacje z modyfikatorami, dopełnieniami i konstrukcjami. Artykuł proponuje sekwencję analizy: ustalanie współżycia, określanie wzorca leksykalno-gramatycznego i interpretacja znaczenia afektywnego. Przykłady głębokiego smutku, przytłaczającego strachu, gorzkiego rozczarowania, gniewu i lęku pokazują, jak dane statystyczne można skorelować z interpretacją filologiczną. Wyniki mogą być wykorzystane w badaniu korpusowym angielskiego słownictwa emocjonalnego oraz w językoznawczej adnotacji językoznawczej (NLP).

Słowa kluczowe: kolokacja, walencja, prozodia semantyczna, preferencja semantyczna, środki asocjacyjne, znaczenie afektywne, konkordancja, adnotacja.

METHODS OF COMBINATORIAL LINGUISTICS IN DESCRIBING EMOTIVE LEXIS: CORPUS-STATISTICAL AND COGNITIVE-SEMANTIC ASPECTS

Nataliia Bober

*Candidate of Philological Sciences, Associate Professor,
Associate Professor at the Department of Germanic Philology,
Faculty of Romance and Germanic Philology,
Borys Grinchenko Kyiv Metropolitan University (Kyiv, Ukraine)*

ORCID ID: 0000-0002-9639-0562

n.bober@kubg.edu.ua

Abstract. The article examines the methods of combinatorial linguistics as a tool for describing emotive lexis. The analysis focuses on collocation, valency, distribution, collostructional relations, semantic preference, semantic prosody and corpus-based association measures. An emotive lexeme is treated not as a self-contained name of an emotion, but as a unit whose meaning is specified through recurrent combinations with modifiers, complements and constructions. The article proposes a sequence of analysis: recording co-occurrence, identifying a lexical-grammatical pattern, and interpreting affective meaning. Examples such as deep sorrow, overwhelming fear, bitter disappointment, anger, and anxiety show how statistical evidence may be correlated with philological interpretation. The results can be used in corpus research on English emotion vocabulary and in linguistically grounded NLP annotation.

Key words: collocation, valency, semantic prosody, semantic preference, association measures, affective meaning, concordance, annotation.

МЕТОДИ КОМБІНАТОРНОЇ ЛІНГВІСТИКИ В ОПИСІ ЕМОТИВНОЇ ЛЕКСИКИ: КОРПУСНО-СТАТИСТИЧНІ Й КОГНІТИВНО-СЕМАНТИЧНІ АСПЕКТИ

Наталія Бобер

*кандидат філологічних наук, доцент,
доцент кафедри германської філології,*

Факультету романо-германської філології Київського столичного університету

імені Бориса Грінченка (Київ, Україна)

ORCID ID: 0000-0002-9639-0562

n.bober@kubg.edu.ua

Анотація. У статті розглянуто методи комбінаторної лінгвістики як інструмент для опису емотивної лексики. У центрі аналізу перебувають колокація, валентність, дистрибуція, колострукційні зв'язки, семантична перевага, семантична просодія та корпусні асоціативні міри. Емотивну лексему трактують не як ізольовану номінацію емоції, а як одиницю, значення якої уточнюється через повторювані поєднання з модифікаторами, комплементами та конструкціями. У статті запропоновано послідовність аналізу: фіксація співживаності, визначення лексико-граматичного патерна та тлумачення афективного значення. На прикладах *deep sorrow*, *overwhelming fear*, *bitter disappointment*, *anger at* і *anxiety about* показано, як статистичні дані можуть бути співвіднесені з філологічною інтерпретацією. Результати можуть бути використані в корпусному дослідженні англійської лексики емоцій та в лінгвістично обґрунтованому NLP-анотуванні.

Ключові слова: колокація, валентність, семантична просодія, семантична перевага, асоціативні міри, афективне значення, конкорданс, анотування.

Introduction. Combinatorial linguistics does not describe a random sum of word combinations. It studies the regular relations through which a word enters speech with a relatively stable set of expected partners. For the vocabulary of emotions, this perspective is especially important. The words *fear*, *anger*, *sadness*, *sorrow*, *joy*, and *anxiety* name emotional states, but their meanings in a text become visible only in the context of other units: *deep sorrow*, *overwhelming fear*, *bitter disappointment*, *anger at injustice*, *anxiety about the future*, and *joy at the news*.

This environment is not secondary material used merely to illustrate a dictionary definition. It shows how the speaker evaluates an emotion, what is taken to be its cause, and how its degree, duration and direction are construed. The adjective 'deep' in 'deep sorrow' does not simply intensify the noun; it introduces a spatial way of conceptualising an inner state. In *overwhelming fear*, the emotion assumes the features of a force beyond the subject's control. In *anger at injustice*, the prepositional complement does not name an accidental addition; it marks the object of evaluative orientation.

The theoretical point of departure is J. R. Firth's well-known claim that a word is known by the company it keeps. In corpus linguistics, this idea was developed by J. Sinclair, M. Stubbs, and M. Hoey, who treat recurrent lexical combinations as evidence for meaning rather than as mere coincidences of words (Firth, 1957; Sinclair, 1991; Stubbs, 2001; Hoey, 2005). For emotive lexis, this entails a shift from lists of emotion names to their combinatorial profiles: collocates, valency frames, semantic classes, evaluative prosody and discourse function.

The aim of the article is to substantiate a philologically workable model for describing emotive lexis by means of combinatorial linguistics. To achieve this aim, the article systematises the methods that can be applied to emotion vocabulary; explains the role of MI-score, t-score and log-likelihood in identifying statistically salient combinations; distinguishes collocation, valency, semantic preference and semantic prosody; and proposes a sequence of analysis that can be reproduced in corpus research on English emotional discourse.

Literature Review. The problem of combinability has several scholarly lines that converge in the analysis of emotive lexis. The first line concerns collocation, phraseology and lexical grammar. Firth's thesis on meaning through habitual context made it possible to regard recurrent combinations as part of a word's semantics (Firth, 1957). Sinclair transferred this thesis into corpus practice by making the concordance and collocation the main tools for describing linguistic regularity (Sinclair, 1991). Stubbs showed that lexical semantics relies on phraseological evidence, while Hoey explained the stability of such evidence through lexical priming – the tendency of a word to occur in particular lexical, grammatical and textual environments (Stubbs, 2001; Hoey, 2005).

For the description of emotive lexis, this line of research leads to a simple but important conclusion: an emotion noun cannot be interpreted in isolation. *Fear* in the frame, *fear of*, behaves differently from sudden *fear*; *anger at injustice* is not identical with suppressed *anger*; sheer *joy* and *joy at the news* have different syntactic and semantic supports. The pattern grammar of S. Hunston and G. Francis is particularly useful here because it allows the researcher to describe not only the word itself, but also the recurrent form through which it enters the sentence (Hunston, Francis, 2000). P. Hanks's work on norms and exploitations in lexical use helps to distinguish a habitual pattern from a stylistically motivated individual use (Hanks, 2013).

The second line concerns the statistical measurement of co-occurrence. K. Church and P. Hanks introduced mutual information into lexicographic analysis of word associations; this measure helps to reveal specific combinations, including those that are not highly frequent (Church, Hanks, 1990: 22–29). T. Dunning proposed log-likelihood as a more reliable tool for rare events and distributions that deviate from expectation (Dunning, 1993:

61–74). F. Smadja described collocation extraction as a sequence of operations: identifying candidate combinations, applying statistical filtering and then proceeding to interpretation (Smadja, 1993: 143–177).

S. Evert demonstrated in detail that no association measure is a universal proof of collocationality: the result depends on frequency, corpus size, span, direction of association and the task for which the measure is used (Evert, 2004; Evert, 2008). P. Pecina and P. Schlesinger showed the advantage of combining several measures in collocation extraction, and Pecina's later work systematised the performance of lexical association measures on a broader empirical basis (Pecina, Schlesinger, 2006: 651–658; Pecina, 2010: 137–158). For corpus comparison, the frequency-profiling method proposed by P. Rayson and R. Garside is also important (Rayson and Garside, 2000: 1–6). S. T. Gries emphasised separately that frequency without dispersion may be misleading: a collocation concentrated in a few texts does not have the same evidential value as a unit distributed evenly across the corpus (Gries, 2008: 403–437; Gries, 2013: 137–166).

The third line links collocation with construction and valency. The collocational analysis developed by A. Stefanowitsch and S. T. Gries provides a tool for describing how a lexical item is attracted to a particular grammatical construction (Stefanowitsch & Gries, 2003: 209–243; Gries & Stefanowitsch, 2004: 97–129). For emotion vocabulary, this is important in constructions such as *be filled with grief*, *burst into tears*, *be overcome by fear* and *be angry at someone*. What is at issue here is not the proximity of two words, but a stable way of grammatically organising an emotional state.

Collocation networks, as proposed by V. Brezina, T. McEnery and S. Wattam, present word meaning through groups of links rather than through a single node-collocate pair (Brezina, McEnery, Wattam, 2015: 139–173). D. Gablasova, V. Brezina and T. McEnery rightly warn that the interpretation of collocational data depends on search parameters, minimum frequency, the selected measure and the aim of the research (Gablasova, Brezina, McEnery, 2017: 155–179). P. Baker, in his analysis of the “shapes of collocation”, shows that collocation should not be reduced to a bigram: in many cases, the grammatical relation, distance between components or wider textual environment is more important (Baker, 2016: 139–164). In practical corpus work, these principles are supported by tools such as Sketch Engine, with its word sketches, grammatical relations and collocational profiles (Kilgarriff et al., 2014: 7–36). General requirements concerning corpus representativeness, annotation, and reproducibility are discussed in the work of C. Manning and H. Schütze, and of T. McEnery and A. Hardie (Manning & Schütze, 1999; McEnery & Hardie, 2012).

The fourth line concerns semantic preference, semantic prosody and evaluation. B. Louw described semantic prosody as the tendency of a unit to acquire evaluative colouring through recurrent environments (Louw, 1993). M. Bednarek clarified that semantic preference refers to the selection of collocates belonging to a particular semantic class, whereas semantic prosody is associated with a positive or negative evaluative background (Bednarek, 2008: 119–139). S. Hunston and A. Partington developed this issue further in corpus-based studies of evaluative language (Hunston, 2011; Partington, 2014). At this point, combinatorial linguistics moves beyond formal counting: it shows how bitter regret, grave concern or sheer delight produce an evaluative image of an emotion.

The fifth line concerns NLP resources for emotion vocabulary. The NRC Emotion Lexicon connects words with emotion categories and polarity; WordNet-Affect presents affective vocabulary as an extension of WordNet; and SemEval-2018 Task 1 proposed an annotation scheme for affect in short digital texts (Mohammad and Turney, 2013: 436–465; Strapparava & Valitutti, 2004; Mohammad et al., 2018: 1–17). SenticNet 7 treats affective meaning within a commonsense-based and neurosymbolic approach (Cambria et al., 2022: 3829–3839). The study by P.-Y. Lu, Y.-Y. Chang and S.-K. Hsieh's analysis of emotion-causing collocations directly confirms that affective meaning often emerges not in a single word but in its combinability (Lu, Chang, Hsieh, 2013: 187–196).

Ukrainian scholarship provides the terminological basis required for such analysis. V. V. Zhukovska systematises the basic concepts of corpus linguistics for philological research (Жуковська, 2013). O. L. Orendarchuk distinguishes combinability and valency within lexical and syntactic combinatorics (Орендарчук, 2012: 258–260). T. V. Bobkova describes Ukrainian corpus lexicography and the methodology for compiling a corpus dictionary of collocations (Бобкова, 2014: 267–272; Бобкова, 2015: 39–41). S. V. Petrasova discusses automatic collocation extraction as a computational-linguistic procedure (Петрасова, 2018: 70–75). O. M. Demska-Kulchytska establishes corpus linguistics as an empirical approach to language in actual use (Демська-Кульчицька, 2005). N. Amosova shows how the concept of collocation may be described through a corpus approach (Амосова, 2013: 60–64).

Materials and Methods. The material of the article consists of the English emotion nouns *fear*, *anger*, *sadness*, *sorrow*, *joy* and *anxiety*, together with typical combinations in which these nouns occur. These units cover different areas of the affective field: *fear*, *anger*, *sadness*, *grief-like sorrow*, *joy* and *anxiety*. They are convenient for methodological description because they differ in valency, collocational range and degree of stylistic markedness. The examples discussed in the article are not presented as final frequency results from a single corpus; they illustrate an analytical order that can be applied to COCA, BNC, NOW, GloWbE, Sketch Engine, or a specially compiled corpus of digital discourse.

The first step is to define the head lexeme. The researcher must properly separate emotive uses from homonymous, terminological, or semantically bleached contexts when such contexts distort the analysis. The second step is to retrieve concordance lines within a fixed span or through grammatical dependencies. For emotive lexis, the grammatical relation is often more informative than mere linear proximity: *fear of death* and *death caused fear* have different syntactic organisation and different semantic focus.

The third step is to group candidate combinations according to structure. Separate groups include adjective + emotion noun (*deep sorrow*, *sudden fear*); verb + emotion noun (*feel fear*, *express anger*); emotion noun +

prepositional complement (anger at N, anxiety about N); predicative adjective + complement (afraid of N, angry with N); and constructions in which an emotion is represented as a cause, force or result (be overcome by fear, burst into tears).

The fourth step is statistical ranking. MI-score helps to identify specific combinations in which the observed frequency is much higher than the expected frequency. It is especially useful for expressive combinations such as utter despair, but it requires caution with low-frequency data. t-score better captures frequency stability and is therefore suitable for combinations such as strong anger, serious concern and growing anxiety. Log-likelihood is used when the researcher needs to test a statistical contrast between genres, registers or subcorpora. In the present article, these measures are not treated as interchangeable: each answers a different question.

The fifth step is philological interpretation. A coefficient does not by itself explain why a combination matters for the description of an emotion. After statistical selection, the researcher must establish whether the collocate functions as an intensifier, an evaluative modifier, a causal indicator, a marker of duration, a marker of suddenness or a means of metaphorical construal. It is at this stage that combinatorial analysis becomes cognitive-semantic analysis.

Results and Discussion. The analysis of the sources and control examples makes it possible to describe an emotive lexeme on at least three levels. The first level is formal co-occurrence. It records that *deep* regularly appears near sorrow, overwhelming near fear, and bitter near disappointment. This level is necessary, but it is not sufficient: word proximity alone does not explain either the function of the modifier or the type of emotional evaluation.

The second level is the lexical-grammatical pattern. *Deep sorrow* belongs to the attributive model adjective + emotion noun; *anger at injustice* realises the model emotion noun + prepositional complement; *be overcome by fear* is a passive construction in which the emotion is construed as a force acting upon the subject. At this level, it becomes clear that affective meaning is distributed across the lexeme, the grammatical frame and the semantics of dependent components.

The third level is cognitive-semantic interpretation. In *deep sorrow*, the adjective *deep* cannot be reduced to a general intensifier. It transfers the spatial schema of depth onto the emotion. The subject does not simply feel sadness; the emotional state is presented as an inner space into which the person is drawn. For this reason, *deep sorrow*, *deep grief*, and profound sadness can be compared not only by frequency but also by how they distribute intensity, style, and imagery.

The combination of *overwhelming fear* works differently. The collocate *overwhelming* presents fear as an excessive force that disrupts the subject's control. In this model, emotion features pressure, dominance, and reduced agency. Corpus analysis should show whether *overwhelming* is attracted mainly to negative affective nouns such as *fear*, *anxiety*, *grief* and *panic*, or whether it also productively combines with positive emotion nouns such as *joy* and *relief*. At this point, the concept of semantic preference becomes a practical analytical tool rather than a decorative term.

In bitter disappointment, another metaphorical basis is at work. Disappointment names a reaction to an unfulfilled expectation, while bitterness transfers a gustatory evaluation onto the experience. This combination does not simply intensify the emotion; it presents it as an unpleasant aftertaste left by an event. If a corpus shows that *bitterness* regularly occurs with disappointment, resentment and regret, the researcher may speak of a preference for negative retrospective emotions. In this case, semantic prosody is formed not by one word but by a series of recurrent evaluative combinations.

Valency analysis provides another type of information. *Fear of*, *anger at*, *anxiety about* and *joy at* should not be treated as ordinary two-component collocations. They are frames that connect an emotion with its object, stimulus or cause. *Fear of* most often opens a position for threat, possible loss or harm. *Anger at* directs the emotion toward an event, person or action that the speaker evaluates as unacceptable. *Anxiety about* profiles an uncertain future, whereas *joy at* profiles a positively evaluated stimulus. Without the description of such frames, the researcher sees the emotion word but not the way it organises the situation.

The near-synonymous nouns *sadness*, *sorrow* and *grief* are especially revealing. In an emotion lexicon, they may fall into the same negative category, but corpus combinatorics separates them. *Grief* predictably gravitates toward loss, death, mourning and ritualised contexts. *Sadness*, more broadly, denotes a psychological state and is more easily used in neutral descriptive contexts. *Sorrow* has a more elevated literary shade and often combines with adjectives of depth and duration. These distinctions cannot be established reliably from a dictionary definition alone; they have to be seen in concordances and collocational series.

The proposed analysis sequence helps avoid three errors. The first is lexical atomism, in which the name of an emotion is taken to be a sufficient carrier of meaning. The second is statistical simplification, in which a high MI-score or log-likelihood value is equated with meaning. The third is intuitive labelling, where a combination is called typical without frequency, association or dispersion evidence. Philological description requires a combination of numerical results, grammatical frames, semantic classes of collocates and contextual interpretation.

To make the procedure suitable for an actual corpus study, the methods should be distinguished according to what each of them observes. Collocational analysis answers the question of which words recurrently co-occur with a head emotive lexeme. Valency analysis shows which positions the lexeme can occupy in a complement. Association measures test whether a combination has statistical weight. Semantic prosody explains the evaluative background created by recurrent environments. Cognitive-semantic analysis clarifies what model of emotion is constructed by the pattern.

Table 1

Combinatorial methods for describing emotive lexis

Method	Unit of observation	Research question	Example
Collocational analysis	recurrent lexical pair or group	which words regularly appear near an emotion noun	deep sorrow; sudden fear
Valency analysis	lexical-grammatical frame	which complements an emotive unit opens	anger at N; anxiety about N
Association measures	observed and expected frequency	whether the combination has statistical weight	MI-score; t-score; log-likelihood
Semantic prosody	recurrent evaluative environment	what evaluative background is created by the collocates	bitter disappointment; grave concern
Cognitive-semantic analysis	conceptual model	how the pattern construes an emotional state	emotion as depth; emotion as force

The table does not replace analysis. It only separates different objects of observation: the word as a head, the combination as a statistical fact, the construction as a grammatical frame, and the pattern as a way of construing an emotion. This distinction prevents the researcher from conflating counting, syntactic description and semantic interpretation.

In empirical work, the researcher may begin with a balanced corpus or with several comparable subcorpora. The list of headwords should be defined by lexical-semantic criteria and, when necessary, checked against the NRC Emotion Lexicon, WordNet-Affect or other affective resources. Concordance lines are then retrieved for each lexeme, and collocates are calculated under the same conditions: the same span, the same minimum frequency threshold and the same set of association measures. If the tool permits it, grammatical dependencies should be used alongside linear proximity.

After automatic selection, manual philological ordering is needed. Combinations should be classified according to structure and function: intensification (deep sorrow, sheer joy), sudden onset of emotion (sudden fear), causal orientation (anger at injustice), future uncertainty (anxiety about the future), the metaphor of force (overwhelming fear) and the metaphor of taste (bitter disappointment). Such classification is not meant to embellish statistics; it explains what the corpus actually measures.

The method can also be applied to Ukrainian material. The nouns *страх*, *гнів*, *радість*, *смуток*, *тривога* and *сум* should not be mechanically equated with their English counterparts. Ukrainian frames such as *страх перед*, *тривога за*, *гнів на* and *радість від*, as well as attributive models such as *глибокий сум*, *щира радість* and *сильний страх*, have their own valency and phraseological restrictions. This is crucial for contrastive studies and for the development of multilingual resources for emotion vocabulary.

The procedure is also suitable for analysing verbal emoticons in digital discourse. If a verbal emoticon is understood as a verbalised marker of affective stance, it can be studied through form, syntagmatic environment, pragmatic function, semantic prosody and annotatability. Emotional meaning in digital text is often not fixed in a single word; it arises from a short but recognisable combination, for example, from an intensifier, an evaluative word, a reactive formula or a graphically marked emotional utterance.

For NLP, this approach has practical value. An emotion lexicon may mark the word 'fear' as an indicator of fear, but it will not always distinguish *irrational fear*, *public fear*, *fear of death*, and *fear for children*. Combinatorial description adds to annotation the information that dictionary labels often lack: the object of emotion, degree of intensity, source of evaluation, type of metaphor, and the speaker's discourse position.

Conclusions. Combinatorial linguistics provides a productive basis for describing emotive lexis because it treats words in terms of their actual relations. For emotion nouns, such relations include collocates, valency complements, grammatical constructions, semantic classes of the surrounding words and evaluative prosody. These relations show how intensity, cause, direction, duration and metaphorical form of an emotional state are shaped in a text.

The article proposes the following sequence of analysis: selecting an emotive lexeme, retrieving concordances, structurally grouping combinations, calculating MI-score, t-score, and log-likelihood, describing the valency frame, identifying semantic preference and prosody, and providing a cognitive-semantic interpretation of the pattern. MI-score is appropriate for finding specific combinations, t-score for frequency-stable models, and log-likelihood for statistically significant contrasts between corpora or genres. None of these measures replaces philological analysis; each only provides an evidential basis for it.

The scholarly value of the proposed approach lies in its description of emotive lexis not as a list of words but as a system of recurrent lexical-grammatical and semantic relations. This model can be used in corpus studies of English and Ukrainian, in contrastive analysis of emotion nouns, in research on digital discourse and in the design of linguistically grounded NLP annotation schemes. A further stage of the research should test this sequence on a balanced corpus and compare the combinatorial profiles of basic emotion nouns in literary, media, academic and digital discourse.

Bibliography:

- Амосова Н. Концепція колокації: корпусний підхід. Науковий вісник Міжнародного гуманітарного університету. Серія: Філологія. 2013. № 10. Т. 2. С. 60–64.
- Baker P. The shapes of collocation. International Journal of Corpus Linguistics. 2016. Vol. 21. No. 2. P. 139–164. <https://doi.org/10.1075/ijcl.21.2.01bak>

3. Bednarek M. Semantic preference and semantic prosody re-examined. *Corpus Linguistics and Linguistic Theory*. 2008. Vol. 4. №. 2. P. 119–139. <https://doi.org/10.1515/CLLT.2008.006>
4. Бобкова Т. В. Українська корпусна лексикографія: основні етапи й тенденції розвитку. Наукові записки Національного університету «Острозька академія». Серія «Філологічна». 2014. Вип. 49. С. 267–272.
5. Бобкова Т. В. Корпусний словник колокацій: методика укладання. Наукові записки Національного університету «Острозька академія». Серія «Філологічна». 2015. Вип. 52. С. 39–41.
6. Brezina V., McEnery T., Wattam S. Collocations in context: a new perspective on collocation networks. *International Journal of Corpus Linguistics*. 2015. Vol. 20. №. 2. P. 139–173. <https://doi.org/10.1075/ijcl.20.2.01bre>
7. Cambria E., Liu Q., Decherchi S., Xing F., Kwok K. SenticNet 7: a commonsense-based neurosymbolic AI framework for explainable sentiment analysis. *Proceedings of LREC 2022. Marseille: ELRA, 2022*. P. 3829–3839.
8. Church K. W., Hanks P. Word association norms, mutual information, and lexicography. *Computational Linguistics*. 1990. Vol. 16. №. 1. P. 22–29.
9. Демська-Кульчицька О. М. Корпусна лінгвістика. Київ: Видавничий дім «Києво-Могилянська академія», 2005. 138 с.
10. Dunning T. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*. 1993. Vol. 19. №. 1. P. 61–74.
11. Evert S. The statistics of word cooccurrences: word pairs and collocations: PhD dissertation. Stuttgart: University of Stuttgart, 2004. 331 p.
12. Evert S. Corpora and collocations. *Corpus Linguistics: An International Handbook* / ed. by A. Lüdeling, M. Kytö. Berlin; New York: De Gruyter, 2008. P. 1212–1248.
13. Firth J. R. *Papers in Linguistics 1934–1951*. London: Oxford University Press, 1957. 233 p.
14. Gablasova D., Brezina V., McEnery T. Collocations in corpus-based language learning research: identifying, comparing, and interpreting the evidence. *Language Learning*. 2017. Vol. 67. S1. P. 155–179. <https://doi.org/10.1111/lang.12225>
15. Gries S. T. Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics*. 2008. Vol. 13. №. 4. P. 403–437. <https://doi.org/10.1075/ijcl.13.4.02gri>
16. Gries S. T. 50-something years of work on collocations: what is or should be next. *International Journal of Corpus Linguistics*. 2013. Vol. 18. №. 1. P. 137–166. <https://doi.org/10.1075/ijcl.18.1.09gri>
17. Gries S. T., Stefanowitsch A. Extending collocation analysis: a corpus-based perspective on alternations. *International Journal of Corpus Linguistics*. 2004. Vol. 9. №. 1. P. 97–129. <https://doi.org/10.1075/ijcl.9.1.06gri>
18. Hanks P. *Lexical Analysis: Norms and Exploitations*. Cambridge, MA: MIT Press, 2013. 480 p.
19. Hoey M. *Lexical Priming: A New Theory of Words and Language*. London; New York: Routledge, 2005. 202 p.
20. Hunston S. *Corpus Approaches to Evaluation: Phraseology and Evaluative Language*. London; New York: Routledge, 2011. 216 p.
21. Hunston S., Francis G. *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam: John Benjamins, 2000. 289 p.
22. Kilgariff A., Baisa V., Bušta J., Jakubíček M., Kovář V., Michelfeit J., Rychlý P., Suchomel V. The Sketch Engine: ten years on. *Lexicography*. 2014. Vol. 1. №. 1. P. 7–36. <https://doi.org/10.1007/s40607-014-0009-9>
23. Louw B. Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. *Text and Technology* / ed. by M. Baker, G. Francis, E. Tognini-24. Bonelli. Amsterdam: John Benjamins, 1993. P. 157–176.
24. Lu P.-Y., Chang Y.-Y., Hsieh S.-K. Causing emotion in collocation: an exploratory data analysis. *Proceedings of PACLIC 27. Taipei, 2013*. P. 187–196.
25. Manning C. D., Schütze H. *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press, 1999. 680 p.
26. McEnery T., Hardie A. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press, 2012. 294 p.
27. Mohammad S. M., Bravo-Marquez F., Salameh M., Kiritchenko S. SemEval-2018 Task 1: Affect in Tweets. *Proceedings of SemEval 2018. New Orleans: ACL, 2018*. P. 1–17. <https://doi.org/10.18653/v1/S18-1001>
28. Mohammad S. M., Turney P. D. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*. 2013. Vol. 29. №. 3. P. 436–465. <https://doi.org/10.1111/j.1467-8640.2012.00460.x>
29. Орендарчук О. Л. Теоретико-методологічні засади вивчення сполучуваності та валентності в аспекті лексичної та синтаксичної комбінаторики. Наукові записки Національного університету «Острозька академія». Серія «Філологічна». 2012. Вип. 26. С. 258–260.
30. Partington A. Evaluative prosody. *Corpus Pragmatics* / ed. by K. Aijmer, C. Rühlemann. Cambridge: Cambridge University Press, 2014. P. 279–303.
31. Pecina P. Lexical association measures and collocation extraction. *Language Resources and Evaluation*. 2010. Vol. 44. №. 1–2. P. 137–158. <https://doi.org/10.1007/s10579-009-9101-4>
32. Pecina P., Schlesinger P. Combining association measures for collocation extraction. *Proceedings of COLING/ACL 2006. Sydney: ACL, 2006*. P. 651–658.
33. Петрасова С. В. Автоматичне видобування колокацій з корпусу текстів. Вісник Національного технічного університету «ХПІ». Серія: Актуальні проблеми розвитку українського суспільства. 2018. № 4. С. 70–75. <https://doi.org/10.20998/2227-6890.2018.4.12>
34. Rayson P., Garside R. Comparing corpora using frequency profiling. *Proceedings of the Workshop on Comparing Corpora. Hong Kong: ACL, 2000*. P. 1–6.
35. Sinclair J. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press, 1991. 179 p.

36. Smadja F. Retrieving collocations from text: Xtract. *Computational Linguistics*. 1993. Vol. 19. №. 1. P. 143–177.
37. Stefanowitsch A., Gries S. T. Collostructions: investigating the interaction of words and constructions. *International Journal of Corpus Linguistics*. 2003. Vol. 8. №. 2. P. 209–243. <https://doi.org/10.1075/ijcl.8.2.03ste>
38. Strapparava C., Valitutti A. WordNet-Affect: an affective extension of WordNet. *Proceedings of LREC 2004*. Lisbon: ELRA, 2004.
39. Stubbs M. *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell, 2001. 267 p.
40. Жуковська В. В. Вступ до корпусної лінгвістики: навчальний посібник. Житомир: Вид-во ЖДУ ім. І. Франка, 2013. 142 с.

References:

1. Amosova, N. (2013). Kontseptsiiia kolokatsii: korpusnyi pidkhiid [The concept of collocation: a corpus approach]. *Naukovyi visnyk Mizhnarodnoho humanitarnoho universytetu. Serii: Filolohiia*, 10(2), 60–64 [in Ukrainian].
2. Baker, P. (2016). The shapes of collocation. *International Journal of Corpus Linguistics*, 21(2), 139–164. <https://doi.org/10.1075/ijcl.21.2.01bak>
3. Bednarek, M. (2008). Semantic preference and semantic prosody re-examined. *Corpus Linguistics and Linguistic Theory*, 4(2), 119–139. <https://doi.org/10.1515/CLLT.2008.006>
4. Bobkova, T. V. (2014). Ukrainska korpusna leksykohefrafii: osnovni etapy y tendentsii rozvytku [Ukrainian corpus lexicography: main stages and development trends]. *Naukovi zapysky Natsionalnoho universytetu «Ostrozka akademii»*. Serii «Filolohichna», 49, 267–272 [in Ukrainian].
5. Bobkova, T. V. (2015). Korpusnyi slovnyk kolokatsii: metodyka ukkladannia [Corpus dictionary of collocations: compilation methodology]. *Naukovi zapysky Natsionalnoho universytetu «Ostrozka akademii»*. Serii «Filolohichna», 52, 39–41 [in Ukrainian].
6. Brezina, V., McEnery, T., & Wattam, S. (2015). Collocations in context: A new perspective on collocation networks. *International Journal of Corpus Linguistics*, 20(2), 139–173. <https://doi.org/10.1075/ijcl.20.2.01bre>
7. Cambria, E., Liu, Q., Decherchi, S., Xing, F., & Kwok, K. (2022). SenticNet 7: A commonsense-based neurosymbolic AI framework for explainable sentiment analysis. In *Proceedings of LREC 2022* (pp. 3829–3839). ELRA.
8. Church, K. W., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics*, 16(1), 22–29.
9. Demska-Kulchyska, O. M. (2005). *Korpusna lingvistika [Corpus linguistics]*. Kyiv: Vydavnychiy dim «Kyievo-Mohylianska akademii» [in Ukrainian].
10. Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1), 61–74.
11. Evert, S. (2004). *The statistics of word cooccurrences: Word pairs and collocations* (PhD dissertation). University of Stuttgart.
12. Evert, S. (2008). Corpora and collocations. In A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics: An International Handbook* (pp. 1212–1248). De Gruyter.
13. Firth, J. R. (1957). *Papers in linguistics 1934–1951*. Oxford University Press.
14. Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, 67(S1), 155–179. <https://doi.org/10.1111/lang.12225>
15. Gries, S. T. (2008). Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics*, 13(4), 403–437. <https://doi.org/10.1075/ijcl.13.4.02gri>
16. Gries, S. T. (2013). 50-something years of work on collocations: What is or should be next. *International Journal of Corpus Linguistics*, 18(1), 137–166. <https://doi.org/10.1075/ijcl.18.1.09gri>
17. Gries, S. T., & Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspective on alternations. *International Journal of Corpus Linguistics*, 9(1), 97–129. <https://doi.org/10.1075/ijcl.9.1.06gri>
18. Hanks, P. (2013). *Lexical analysis: Norms and exploitations*. MIT Press.
19. Hoey, M. (2005). *Lexical priming: A new theory of words and language*. Routledge.
20. Hunston, S. (2011). *Corpus approaches to evaluation: Phraseology and evaluative language*. Routledge.
21. Hunston, S., & Francis, G. (2000). *Pattern grammar: A corpus-driven approach to the lexical grammar of English*. John Benjamins.
22. Kilgariff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, 1(1), 7–36. <https://doi.org/10.1007/s40607-014-0009-9>
23. Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and Technology* (pp. 157–176). John Benjamins.
24. Lu, P.-Y., Chang, Y.-Y., & Hsieh, S.-K. (2013). Causing emotion in collocation: An exploratory data analysis. In *Proceedings of PACLIC 27* (pp. 187–196).
25. Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. MIT Press.
26. McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.
27. Mohammad, S. M., Bravo-Marquez, F., Salameh, M., & Kiritchenko, S. (2018). SemEval-2018 Task 1: Affect in Tweets. In *Proceedings of SemEval 2018* (pp. 1–17). ACL. <https://doi.org/10.18653/v1/S18-1001>
28. Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3), 436–465. <https://doi.org/10.1111/j.1467-8640.2012.00460.x>
29. Orendarchuk, O. L. (2012). *Teoretyko-metodolohichni zasady vyvchennia spoluchuvanosti ta valentnosti v aspekti leksychnoi ta syntaksyschnoi kombinatoryky [Theoretical and methodological principles of studying combinability and*

- valency in lexical and syntactic combinatorics]. Naukovi zapysky Natsionalnoho universytetu «Ostrozka akademiia». Seriiia «Filolohichna», 26, 258–260 [in Ukrainian].
30. Partington, A. (2014). Evaluative prosody. In K. Aijmer & C. Rühlemann (Eds.), *Corpus Pragmatics* (pp. 279–303). Cambridge University Press.
 31. Pecina, P. (2010). Lexical association measures and collocation extraction. *Language Resources and Evaluation*, 44(1–2), 137–158. <https://doi.org/10.1007/s10579-009-9101-4>
 32. Pecina, P., & Schlesinger, P. (2006). Combining association measures for collocation extraction. In *Proceedings of COLING/ACL 2006* (pp. 651–658). ACL.
 33. Petrasova, S. V. (2018). Avtomatychnе vydobuvannia kolokatsii z korpusu tekstiv [Automatic extraction of collocations from a text corpus]. *Visnyk Natsionalnoho tekhnichnoho universytetu «KhPI»*. Seriiia: Aktualni problemy rozvytku ukrainskoho suspilstva, 4, 70–75. <https://doi.org/10.20998/2227-6890.2018.4.12> [in Ukrainian].
 34. Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. In *Proceedings of the Workshop on Comparing Corpora* (pp. 1–6). ACL.
 35. Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
 36. Smadja, F. (1993). Retrieving collocations from text: Xtract. *Computational Linguistics*, 19(1), 143–177.
 37. Stefanowitsch, A., & Gries, S. T. (2003). Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics*, 8(2), 209–243. <https://doi.org/10.1075/ijcl.8.2.03ste>
 38. Strapparava, C., & Valitutti, A. (2004). WordNet-Affect: An affective extension of WordNet. In *Proceedings of LREC 2004*. ELRA.
 39. Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Blackwell.
 40. Zhukovska, V. V. (2013). Vstup do korpusnoi linhvistyky [Introduction to corpus linguistics]. Zhytomyr: Vyd-vo ZhDU im. I. Franka [in Ukrainian].